Heterogeneity of innovative, collaborative, and productive firm-level processes

Amoroso, S.

Heterogeneity of innovative, collaborative, and productive firm-level processes

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op woensdag 30 januari 2013 om 10.15 uur door

SARA AMOROSO,

geboren op 7 januari 1981 te Rome, Italië.

Promotores:	prof.dr. Bertrand Melenberg
	prof.dr. Joseph E. Plasmans
Overige leden van de Promotiecommissie:	prof.dr. Peter M. Kort
	prof.dr. Dirk Czarnitzki
	dr. Sabien Dobbelaere
	dr. Tobias J. Klein

Acknowledgements

The past four years have been a sequence of unforgettable events, and not everything has always been smooth sailing. Nevertheless, if it were not for the support of many people, I would have never become conscious of the fact that “there is a crack in everything, that’s how the light gets in” (Leonard Cohen, *Anthem*).

I owe my deepest gratitude to my supervisors, Joseph Plasmans and Bertrand Melenberg, for their continuous advice and encouragement throughout the course of my doctorate. Jef transmitted to me his enthusiasm and passion for many relevant topics in the field of Econometrics of Innovation. Bertrand represented a moral and academic example to live up to. From the inception to the completion of this thesis, he devoted himself to my study, challenging me to look into new topics from different perspectives.

I am honored to have dr. Tobias Klein, prof. dr. Peter Kort, prof. dr. Dirk Czarnitzki, dr. Sabien Dobbelaere as Ph.D. committee members. I am indebted for their helpful comments during my pre-defense.

I would also like to express my very sincere gratitude to my MSc thesis supervisor, prof.dr. Brunero Liseo, as he always provided me with help and support during my studying here in Tilburg.

My life in Antwerp and in Tilburg would have been all the less enjoyable were it not for the support and caring provided by all my colleagues and friends. Thanks to all of you. A special thanks goes to Raissa, Yan, Gaia, Juan, Novella, and Mara. They say it is hard to find a friend...then I must be lucky, for I have found six!

Finally, I acknowledge my gratitude to my family, whose unconditional love and support encouraged me to go ahead and take chances in life.

Sara Amoroso
Tilburg, January 2013

Contents

Chapter 1

Introduction

During the last forty years, economics of innovation has emerged as a distinct area of enquiry at the crossing of the economics of growth, industrial organization, regional economics, and the theory of the firm. It became a well identified area of competence in economics specializing not only in the analysis of the effects of the introduction of new technologies, but also and mainly in understanding technological change as an endogenous process. As the result of the interpretation, elaboration, and evolution of different fields of analysis in economic theory, innovation is viewed as a complex, path dependent process characterized by the interdependence and interaction of a variety of heterogeneous agents.

Several features of a complex industrial innovation system deserve greater attention in economics. First, the firms are characterized by distinctive and specific characteristics as well as being intrinsically heterogeneous. Second, there are many levels of organization and interaction in an economy. Therefore, behaviors, actions, strategies, products at one level typically serve as "building blocks" for constructing units at the next higher level. The overall organization is hierarchical, with many sorts of tangling interactions (associations, channels of communication) across levels.

This dissertation presents a structural analysis of the heterogeneity of innovative, collaborative, and productive behaviors of firms. Both innovation and production processes, as well as the formation of R&D collaborations, constitute a challenge for the researcher, as they are interconnected black boxes: labor, materials, and capital, or R&D expenditure and knowledge spillovers entering as inputs from one side, and (innovative) sales, or cooperation strategies, as output resulting from the other. In this context, formulating the accurate measure of productivity is key to assess how efficiently a firm is turning inputs into outputs. Furthermore, the level of efficiency of a firm is affected by a complex set of factors: absorptive capacity, innovativeness, investments in R&D, exporting decisions, and so forth. These factors are, in turn, determined by research networks, labor skills, and the social embeddedness of firms routines and strategies.

The main body of the thesis consists of three empirical essays.

Chapter 2 investigates the potential benefits associated with the presence of unions in terms of productivity growth. Theoretically, the presence of trade unions is arising from the asymmetry in contracting between individual workers and employers. However, the absence of unions may not correspond to an underlying perfectly competitive situation in the labor market. Instead, it may lead to market imperfections on the labor demand side in the form of monopsony, that is, a situation in which there is only one buyer of the labor services. Additionally, firms in some industries may pay workers more than the going market rate to attract new ones. Hence, policy makers, whose objective is to establish and maintain a perfectly competitive labor market, seek for policies designed to free up the demand side of the market. The presence of unions in such circumstances may offer a second-best alternative to free competition. Moreover, the potential benefits associated with the presence of unions in the form of "voice" through which the employees can express their grievances without having to leave the firm , should be counted against the costs due to misallocation effects. As a matter of fact, unions can contribute positively to the productivity of a firm, by facilitating the communication between labor force and management, drawing the attention of the latter to changes in working methods or production techniques that may be beneficial to both parties.

There exist abundant theoretical arguments for both positive and negative effects of trade unions on productivity, but data appropriate for investigating the question are generally lacking. In particular, data for union versus nonunion labor are not available, and even unionization data are highly aggregated. Therefore, we propose a methodology to evaluate the impact of unions on total factor productivity (TFP) growth that does not rely on individual worker data, but on firm-level data. In particular, we rely on the framework of production function to estimate simultaneously price-cost markups and wage markups.

In **Chapter 3**, we analyze the correlation among different R&D alliances due to both firm- and sector-level heterogeneity, adopting a multivariate hierarchical logit model. The advantage of adopting a hierarchical structure (often referred to as multilevel, random or mixed) is that we can control for a richer variance structure that describes the expected correlation present among firms' cooperative choices within a particular sector. In fact, when exploring the differences in the factors affecting the firms' probability to establish different types of cooperation, namely horizontal cooperation (with competitors), vertical cooperation (with suppliers or customers), and institutional cooperation (with universities and research institutes), the existing literature on the determinants of R&D cooperation with different R&D partners overlooks the role of sectoral specificities (for example, sector-specific physical assets) in influencing the expected correlation among the different cooperative strategies present among firms within a particular sector. As a matter of fact, horizontal R&D cooperations are likely to be formed within the same sector as it will lead to *collective efficiencies* (Schmitz, 1999) in the form of reduced transaction costs and accelerated innovation rates through a greater market access. These collective efficiencies are of particular interest from a policy perspective. In fact, they can be interpreted as the multiplier effect of an innovation policy: the increased innovative intensity of one company or several companies multiplies the economic benefits in a given sector by helping to drive the innovative intensity of other business entities. This type of externality is demand-driven, in the sense that the private and public for innovation can be stimulated by new or growing business enterprises, which enables their suppliers to grow as well.

Our approach departs from the one used to test for complementarities (Mohnen and Roller, 2005; Belderbos et al., 2006), as our main focus is not to estimate the degree of strategic complementarities or substitutabilities among firms' cooperation choices, but rather to model and estimate both individual and aggregate forms of externalities, in which the collective actions of a reference group affect an individual's choices. As pointed out by Mohnen and Roller (2005), innovation policies may have different impacts on the distinctive phases of the innovation process. As a matter of fact, while there could be firm-level policy externalities in the decision to collaborate in research, the innovative produce might well be affected by the aforementioned demand driven innovation policy externality.

Therefore, to explicitly take into account both firm- and sector-level externalities, and the different impact of innovation policy measures on the two phases of the innovative process, we divide our study in two stage. In the first stage we study the main drivers of undertaking a collaborative agreement with a research partner. In the second stage, we investigate the effects of innovation policies and R&D cooperation on innovative intensity.

Our hypothesis of a heterogeneity across firms and sectors is confirmed by the results. All covariances are found to be significant. In other words, firms within the same industry share similar characteristics (same random effects), which lead to correlation between research partners' choices.

In **Chapter 4**, we construct a model where firms invest in R&D activities with or without a research partner to improve their productivity levels. In particular, we develop and estimate a structural dynamic monopoly model to quantify the linkages between R&D spending, innovation and cooperation investment choices, and endogenous productivity. To our knowledge, our paper constitutes the first attempt to explicitly model the different collaborative R&D investment decisions adopting a dynamic structural framework. Differently from these studies, the model we propose derives the firms' optimal R&D investment decisions where these depend on the past R&D activities and on the past level of productivity. Additionally, within the suggested framework we are able to model and retrieve the current fixed or sunk costs relative to the different (collaborative) R&D activities.

Merging data on sales and factor inputs of Dutch manufacturing firms extracted from the Production Survey (PS), and three waves of the Community Innovation Survey (CIS) for the Netherlands, covering the period from 2002 to 2008, we find that the firm's probability to do R&D or to introduce an innovation increases with the level of productivity, only when this activity is shared with a research partner. Moreover, past investments in cooperative research have a positive impact on current productivity, which, in turn, positively influences the probability to engage in both R&D and innovation when these activities are shared with a research partner. Therefore, innovation policy measures targeting cooperation might result in a virtuous cycle.

Moreover, according to the literature on R&D cooperation, the costs of innovating are smaller when cooperating. In fact, given the higher risks associated with the uncertainty of the market demand for new products or processes, the firm might allocate more importance to the cost/risk sharing rationale for this type of innovative activities, rather than for the sheer research investments.

Chapter 2

Firm-level productivity under imperfect competition in output and labor markets

Joint work with B. Melenberg, J.E. Plasmans, and M. Vancauteren

Abstract. This paper examines the interaction between product and labor market imperfections and the impact of these imperfections on total factor productivity (TFP) growth. We contribute to the existing literature in several ways. First, we link the empirical research on production functions and the market imperfections literature, providing empirical evidence of the imperfect competition on both product and labor markets. In particular, we present a new way to model how firms deal with output and labor rigidities, and, at the same time, we address potential endogeneity issues including measurement errors in output, and highly correlated factor inputs. We review up–to–date estimation approaches, and confront the validity of the different assumptions in identifying the production function parameters. Second, we consider the role of adjustment frictions in materials in solving the identification issue due the collinearity among productive inputs. Third, using a firm-level dataset of 21 Dutch manufacturing industries over the period 1989–2008, we consider to what extent the estimated productivity is sensitive to the different model specifications and to the different econometric approaches.

2.1 Introduction

Theoretically, the presence of trade unions is arising from the asymmetry in contracting between individual workers and employers. The alternative to a unionized labor market is one characterized by a perfectly competitive structure that ensures the workers to choose whether or not to work by comparing the given perfectly competitive wage with the marginal utility of not working. However, the absence of unions may not correspond to an underlying perfectly competitive situation in the labor market. Instead, it may lead to market imperfections on the labor demand side in the form of monopsony, that is, a situation in which there is only one buyer of the labor services. Additionally, firms in some industries may pay workers more than the going market rate to attract new ones. Hence, policy makers, whose objective is to establish and maintain a perfectly competitive labor market, seek for policies designed to free up the demand side of the market. The presence of unions in such circumstances may offer a second-best alternative to free competition. Moreover, the potential benefits associated with the presence of unions in the form of "voice"[1] should be counted against the costs due to misallocation effects. As a matter of fact, unions can contribute positively to the productivity of a firm, by facilitating the communication between labor force and management, drawing the attention of the latter to changes in working methods or production techniques that may be beneficial to both parties.

There exist abundant theoretical arguments for both positive and negative effects of trade unions on productivity, but data appropriate for investigating the question are generally lacking. In particular, data for union versus nonunion labor are not available, and even unionization data are highly aggregated. Therefore, in this paper, we propose a methodology to evaluate the impact of unions on productivity growth that does not rely on individual worker data, but on firm-level data. In particular, we rely on the framework of production function to estimate simultaneously price-cost markups and wage markups. To our knowledge, this study is the first attempt of estimating simultaneously wage and product markups, and productivity growth, adopting a production function framework.

[1] A source of empowerment through which the employees can express their grievances without having to leave the firm (Aidt and Tzannatos, 2008)

Within the production function framework, the econometric literature has focused on some important econometric issues that need to be dealt with when estimating such functions at the firm level. A first econometric issue, broadly known as simultaneity, is the potential correlation between unobserved productivity shocks and the input factors. A second problem arises from measurement errors in output or inputs. Typically, we observe deflated measures in place of the original physical quantities. This leads to two kinds of problems: the first is associated with endogeneity and the second concerns the correct identification of firm–specific productivity measures. The endogeneity problem is due to the potential correlation between unobservable firm-level (input and output) price variations and input choices (Klette and Griliches, 1996; Ornaghi, 2008; De Loecker, 2011). The productivity measurement identification issue is deriving from the fact that neglecting the variation in factor prices leads to the estimation of performance measures that are indices of revenue per unit input expenditure, rather than measures of efficiency (Katayama et al., 2009). Another problem concerns the highly correlated factor inputs. In fact, if inputs are chosen optimally with no adjustment costs or frictions, the input levels are perfectly dependent of each other (the "collinearity problem"). This collinearity issue may pose a problem especially for some estimators which rely on a control function approach of production function parameters (see Gandhi et al. (2011), Wooldridge (2009), and Ackerberg et al. (2006)'s critique on both Olley and Pakes (1996) and Levinsohn and Petrin (2003)). Lastly, some specific functional form assumptions can be more appropriate than others, due to, for example, the particular type of labor market.

Parallel to the econometric literature, two other strands of research focus on market imperfections. One is the literature on output market imperfect competition which follows the lead of the seminal papers of Hall (1986, 1988, 1991). The other strand, led by McDonald and Solow (1981), focuses on imperfections in labor markets due to unionization.

Only a few studies empirically investigate the possibility of having imperfect competition in both product and labor factor markets. Among others, Bughin (1993, 1996), Crépon et al. (2002), Dobbelaere (2004), Galí et al. (2007), Abraham et al. (2009), and Dobbelaere and Mairesse (2011), consider the possibility of imperfections in both product and factor markets, by taking into account the fact that wages are no longer exogenous.

Bughin (1993), studying the Belgian chemical industry (and four Belgian manufacturing sectors, Bughin (1996)), considers imperfections in product and factor markets, but does not provide insights about the unobserved productivity. Moreover, he does not consider possible endogeneity issues tailored to the selection bias (due to the omission of firms' entry and exit). Also in Crépon et al. (2002), Dobbelaere (2004), and in Dobbelaere and Mairesse (2011), the main focus is the heterogeneity in price–cost markup and workers bargaining power parameters, rather than on productivity and endogeneity issues. Abraham et al. (2009), using Belgian firm–level data, simultaneously estimate price–cost margins and union bargaining power to analyze how price setting and bargaining power is affected by globalization. Although Abraham et al. (2009) apply the Olley and Pakes (1996) method to deal with the simultaneity issue, they do not correct for the unavailability of physical output volumes, replacing the volumes of productions with deflated firm-level sales. The omission of the output price might lead to correlation between the input choices and the productivity shock, yielding biased estimated coefficients.

Next to the main contribution of assessing the impact of unions' power on productivity growth, this paper bridges the gap between the empirical research on production functions and the market imperfections literature. We provide a way to model how firms deal with output and labor market rigidities, and, at the same time, we address the potential endogeneity issues concerning the simultaneity, the collinearity of inputs, and the omitted output price biases. Second, we consider to what extent the estimated unobserved productivity is sensitive to the different model specifications and to the different econometric approaches to identify the structural parameters of the model. Third, we provide further empirical evidence of the imperfect competition on the product and labor markets using up–to–date estimation approaches.

In particular, using a firm-level dataset of 21 Dutch manufacturing industries over the period 1989–2008, we show that, neglecting a *wage markup*, therefore assuming that firms are setting the marginal revenue product of labor equal to the labor's marginal cost, leads to an underestimation of the true value of the price–cost margin at the aggregate level. Depending on the estimation approach, the underestimation of the product markups is significant and varies between 7% and 16%. The workers' bargaining power parameters range between 0.191 and 0.482. These results, except for the within estimator, are consistent with the trade union density/unemployment rate ratio reported for the Netherlands. Along with the bargaining parameter, we are able to provide, within the deflated revenue production function framework, an estimate of the *wage markup*. The wage markup estimates fluctuates between 21.9% and 23.7%. These results are comparable with Aidt and Tzannatos (2008)'s review of wage markups in high–income economies, which are between 0 and 25%.

In line with the literature, we find that omitting output prices yields downward biased input elasticity coefficients. We then confirm the hypothesis of sectoral specificity as suggested by Ramey and Shapiro (2001), namely, sectoral specificity concerning physical capital, which is costly to redeploy, sectoral nature of the local labor markets, workforce training institutions, and financial institutions. Indeed, testing the hypothesis of heterogeneity across sectors yields the conclusion that all the structural parameters significantly differ from sector to sector, and are sensitive to the estimation technique.

Moreover, we find that the underestimation of the product markups derives from a computational bias of the true level of the output markup, possibly caused by the misspecification of the marginal costs, as we are omitting the direct effects of wage rigidities. Therefore, firms share their monopoly rents with labor unions.

Concerning the impact of different estimation approaches and model assumptions on the firm-level productivity, with each estimation approach, we find evidence of a positive time trend only when we consider both labor and product market imperfections.

We find that correcting for the omitted output prices leads to decreases of the TFPG percentage rates, and, independently from the estimation technique used, when assuming both bargaining on the labor market and imperfect competition on the output market, we find larger TFPG rates. With both IV and Levinsohn and Petrin (2003) estimation of the TFP growth we find a significant and positive relation between productivity growth and the bargaining parameter. The economic benefits of unions could be found in the worker–manager cooperation.

Indeed, unions can increase firms' productivity by "shocking" the management into better production practices (Aidt and Tzannatos, 2008).

The remainder of the paper is organized as follows. In Section 2, we formulate a measure of total factor productivity (TFP) that allows for both output and labor market power. Section 3 reviews the main estimation techniques. Section 4 describes the data and results on the relevant structural parameters are reported in Section 5. In Section 6 we discuss the results for the TFP measure. In the final section we conclude.

2.2 The model: Building blocks

2.2.1 The standard setting

As in Ackerberg et al. (2007), the starting point is a Cobb-Douglas production function, where the gross output Q_{it} of firm i at time t relates to three specific inputs as follows:

$$Q_{it} = A_{it} K_{it}^{\theta_{iKt}} L_{it}^{\theta_{iLt}} M_{it}^{\theta_{iMt}}, \tag{2.1}$$

where K_{it} denotes capital, L_{it} labor, and M_{it} intermediate goods, consisting of materials and energy, for firm i at period t. A_{it} represents the Hicksian neutral efficiency level of firm i at time t, and is defined as Total Factor Productivity[2](TFP). $\theta_{iKt}, \theta_{iLt}, \theta_{iMt}$ are firms' elasticities of output with respect to capital, labor, and intermediate goods, respectively. Taking natural logs of (4.2) results in a linear function,

$$q_{it} = \theta_0 + \theta_{iKt} k_{it} + \theta_{iLt} l_{it} + \theta_{iMt} m_{it} + a_{it} \tag{2.2}$$

where lower-case letters refer to natural logarithms. The logarithm of A_{it} is defined as $\log(A_{it}) \equiv \theta_0 + a_{it}$, where θ_0 measures the mean productivity level across firms and over time, while a_{it} is the productivity shock which is observable by

[2]MFP (*Multi-Factor Productivity*) is sometimes used interchangeably with TFP, even if there is a slight difference between what they may include. Indeed, taking into account all the factors influencing output levels can be unrealistic, therefore MFP may be a more appropriate term to use. However, the term TFP continues to be used more widely.

the firm (for example, managerial ability, quality of research), but unobservable to the econometrician, hence a source of potential endogeneity.

The time-varying, input-dependent elasticity of scale θ_{it} is defined as the sum of all output elasticities with respect to the three nonnegative factor inputs, X_{ikt}:

$$\theta_{it} \equiv \sum_{k \in \{K,L,M\}} \frac{\partial Q_{it}}{\partial X_{ikt}} \frac{X_{ikt}}{Q_{it}} \equiv \sum_{k \in \{K,L,M\}} \theta_{ikt}.$$

2.2.2 Imperfect output market competition: considering the omitted prices

As we observe deflated gross output, input coefficients might be biased if firm-level price variation is correlated with input choice. To see this, we can express the deflated gross output as $Y_{it} \equiv \frac{Q_{it}(P_{it})P_{it}}{P_t^j} \exp(u_{it}^y)$, where P_{it} is the price of firm i at time t, P_t^j is the industry j ($\equiv j(i)$) price index, and u_{it}^y represents measurement error in Y_{it}. In logs, we have:

$$y_{it} = q_{it} + (p_{it} - p_t^j) + u_{it}^y. \tag{2.3}$$

Substituting equation (2.2) into (2.3), and taking y_{it} as dependent variable, the unobserved firm-level price deviations $(p_{it} - p_t^j)$ will enter the production function as an extra error component. This will introduce potential correlation with the input choices, if $E(x_{it}(p_{it} - p_t^j)) \neq 0$, where $x_{it} \equiv (l_{it}, k_{it}, m_{it})'$, possibly yielding biased input coefficients (Klette and Griliches, 1996; De Loecker, 2011). In order to estimate the production function consistently, without information on establishment-level prices, we proceed by imposing some structure on the demand system (Foster et al., 2008).

Following Klette and Griliches (1996) and De Loecker (2011), a simple conditional Dixit-Stiglitz demand system is expressed as:

$$Q_{it} = Q_t^j (P_{it}/P_t^j)^{\eta_j} \exp(u_{it}^d) \Leftrightarrow P_{it} = P_t^j (Q_{it}/Q_t^j)^{1/\eta_j} \exp(-u_{it}^d/\eta_j) \tag{2.4}$$

where Q_t^j is the sector j production index,[3] and u_{it}^d is an idiosyncratic firm-specific demand shock. Moreover, assuming that consumers have an unbounded taste for variety, it is reasonable to assume that every firm will produce a distinct variety, and that the aggregate production index is, and η_j is the price elasticity of demand for differentiated goods in sector j. Taking logarithms, we can write the right hand side of equation (4.1) as the inverse demand function:

$$p_{it} = p_t^j + \frac{1}{\eta_j}(q_{it} - q_t^j) - \frac{1}{\eta_j}u_{it}^d. \tag{2.5}$$

Taking into account the demand (2.5), the log deflated output (2.3) can be expressed as

$$y_{it} = q_{it} + \frac{1}{\eta_j}(q_{it} - q_t^j) - \frac{1}{\eta_j}u_{it}^d + u_{it}^y. \tag{2.6}$$

Finally, combining equations (2.2) and (2.6), and defining the price *markup* as $\mu_j \equiv \frac{1}{1+1/\eta_j} = \eta_j/(1+\eta_j)$,[4] where $\eta_j < -1$, the deflated gross output can be written as:

$$y_{it} = \gamma_{i0t} + \gamma_{iKt}k_{it} + \gamma_{iMt}m_{it} + \gamma_{iLt}l_{it} - \frac{1}{\eta_j}q_t^j + \tilde{a}_{it} + \tilde{u}_{it}, \tag{2.7}$$

where $\gamma_{i0t} \equiv \theta_0/\mu_j, \gamma_{ikt} \equiv \theta_{ikt}/\mu_j, k = K, L, M$ are the input factor elasticities, $\tilde{a}_{it} \equiv a_{it}/\mu_j$ the productivity shock, $\tilde{u}_{it} \equiv \tilde{u}_{it}^d + \widetilde{u}_{it}^y$, where $\tilde{u}_{it}^d \equiv -u_{it}^d/_j$ is the demand shock and $\widetilde{u}_{it}^y$ is the measurement error in y_{it}.

By introducing a firm's demand system when modeling a production function, we are able to decompose the traditional measured productivity gains into real productivity gains, a_{it}, an idiosyncratic firm–specific demand shock, u_{it}^d and a measurement error, u_{it}^y. Moreover, as in Klette and Griliches (1996), we derive an expression for deflated revenue in an imperfectly competitive framework, where it is possible to obtain an estimate of the demand elasticity η_j, by simply adding industry output as an additional regressor to proxy for unobserved firm–level

[3] We assume the aggregate production index to be exogenous and independent from the output produced by firm i, in other words, $\partial Q_t^j/\partial Q_{it} = 0, \forall t$.

[4] In case of perfect competition, the cross price elasticity tends to minus infinity and the markup goes to one.

prices.[5] The advantage of this structural approach consists in estimating the production function coefficients, controlling for price and demand variation. In this way, we get rid of the potential correlation between measured productivity and all those factors that might have an impact on prices and demand, but are not related to the true productivity (for example, in open economies, real exchange rate appreciation pulls down the output prices of the tradeable goods).

2.2.3 Labor market rigidities: union bargaining power

In this section, we relax the conventional assumption of perfect competition in the labor market, allowing both firms and workers' union to have some market power.

Many authors have studied the influence of market power of unions, by introducing wage rigidities through efficiency wages. For instance, Hall (1991), following McDonald and Solow (1981), assumes that the firm wages and level of employment are jointly determined according to an efficient bargaining scheme between the firm and its workers. In this case, the wage of workers is determined at a level which is higher than the firm's marginal revenue of labor. Workers in firms with some degree of market power on the output market can earn wages that are much higher than the competitive industry wage level.

The workers in the firm bargain with the firm over both the levels of employment and of the wage. In particular, the workers' objective is specified as the union's aggregate gain to the workers from membership, $U_{it}(W_{it}, L_{it}) \equiv L_{it}(W_{it} - \overline{W}_{it})$, where $\overline{W}_{it}$ is the reservation wage (i.e., the theoretical wage valid on an imperfectly competitive output market and a perfectly competitive labor market), and

[5] This model has been first proposed by Klette and Griliches (1996), but De Loecker (2011) was the first to implement correction for output market imperfection into the semi–parametric estimation framework introduced by Olley and Pakes (1996).

W_{it} is the negotiated wage.[6] The standard static firm's objective is to maximize its short-run profit, Π_{it}, given by the difference between the total revenue and the total costs, $\Pi_{it} \equiv P_{it}(Q_{it})Q_{it} - W_{it}L_{it} - R_{it}K_{it} - Z_{it}M_{it}$. We assume that there are adjustment frictions in both the capital and material markets, where the capital and material input levels are held fixed for a time period, and the maximization of Π_{it} with respect to K_{it} and to M_{it} yields the trivial solution of marginal product of such inputs equal to zero.

The efficient bargaining model can be written as a weighted average of the logarithms of workers' aggregate gain from union membership and the firm's profit:

$$\max_{W_{it},L_{it},K_{it},M_{it}} \left[\phi_{it} \log(U_{it}(W_{it}, L_{it})) + (1-\phi_{it}) \log \Pi_{it}\right],$$

where $\phi_{it} \in [0,1]$ is the degree of union bargaining power. The maximization of the union is with respect to the negotiated wage W_{it} and labor L_{it}, but also possibly with respect to the factor inputs K_{it} and M_{it}, depending on whether one asumes these to be flexible or not. The (relevant part of the) FOCs of this problem can be written as:

$$w.r.t. \quad L_{it} \quad \rightarrow \quad (1-\phi_{it})\frac{W_{it} - \left(1+\frac{1}{\eta_j}\right) P_{it}(Q_{it})\frac{\partial Q_{it}}{\partial L_{it}}}{\Pi_{it}} = \frac{\phi_{it}}{L_{it}}, \tag{2.8}$$

$$w.r.t. \quad W_{it} \quad \rightarrow \quad (1-\phi_{it})\frac{W_{it} - \overline{W}_{it}}{\Pi_{it}} = \frac{\phi_{it}}{L_{it}}. \tag{2.9}$$

From (4.21) it follows that, when $\phi_{it} = 0$, the marginal revenue product of labor is equal to the wage rate. However, in general, with $\phi_{it} \in (0,1)$, by rewriting equation (4.22), we can express the bargained wage rate as a function of the

[6] According to McDonald and Solow (1981) the workers' objective in their efficient bargaining model can be specified in two alternative ways: either as the union's aggregate gain to the workers from membership, or taking account of the unemployment benefits, as $L_{it}W_{it} + \overline{W}_{it}(N_{it} - L_{it})$, where N_{it} is the labor supply. McDonald and Solow (1981) judge the first specification as the most appropriate one for real life. In fact, in the second specification, if $\overline{W}_{it}$ falls, the firm would have to increase its wage offer to make up for a reduction in $\overline{W}_{it}$, to keep the level of union utility unchanged. Hence, we advocate McDonald and Solow (1981)'s suggestion and take the union preferences $U_{it}(W_{it}, L_{it}) \equiv L_{it}(W_{it} - \overline{W}_{it})$ as a function of both wages and employment.

bargaining parameter, ϕ_{it}, and the ratio between profits and cost of labor:

$$\frac{W_{it} - \overline{W}_{it}}{W_{it}} = \frac{\phi_{it}}{1-\phi_{it}} \frac{\Pi_{it}}{L_{it} W_{it}}. \tag{2.10}$$

Defining $\mu_{it}^{W} \equiv \frac{W_{it}-\overline{W}_{it}}{W_{it}}$ as the *wage markup*, one can see how this is directly depending on the union's bargaining power. Equation (2.10) summarizes the features of the efficient bargaining model. The wage wedge $W_{it}-\overline{W}_{it}$ is increasing with the bargaining power ϕ_{it} and with firm performance, measured as profit per employee Π_{it}/L_{it}. The profit per employee is a good measure of firm performance, as it proxies for earnings on intangible assets, such as knowledge, reputation, and collaboration created by talented workers (Bryan, 2007).

Combining equations (4.21) and (4.22), we can write the marginal revenue product of labor as

$$\left(\frac{\eta_j + 1}{\eta_j}\right) P_{it}(Q_{it}) \frac{\partial Q_{it}}{\partial L_{it}} = \overline{W}_{it}. \tag{2.11}$$

Therefore, by multiplying both sides of (4.23) by $\frac{L_{it}}{Q_{it}}$, and using the definitions of θ_{iLt} and μ_j, we can derive

$$\gamma_{iLt} \equiv \theta_{iLt}/\mu_j = \frac{\overline{W}_{it} L_{it}}{P_{it}(Q_{it}) Q_{it}} = \frac{\overline{W}_{it}}{W_{it}} \frac{W_{it} L_{it}}{P_{it}(Q_{it}) Q_{it}} \equiv \frac{\overline{W}_{it}}{W_{it}} \times s_{iLt},$$

with s_{iLt} the labor share in total production. Thus, under imperfect competition in both output and labor market, the labor elasticity is a function of the labor share and the wage markup:

$$\gamma_{iLt} = s_{iLt}(1 - \mu_{it}^{W}). \tag{2.12}$$

At this stage, it is intuitively clear how the exclusion of frictions in the labor market (i.e., $\phi_{it} = 0$ or $W_{it} = \overline{W}_{it}$) might lead to misestimating the firm's market power. When there is no imperfect competition in the labor market, firms set the wage at the lowest value possible, ultimately equal to the competitive wage, i.e., $W_{it} = \overline{W}_{it}$ (and, therefore, $\mu_{it}^{W} = 0$). For W_{it} that tends to $\overline{W}_{it}$, the wage markup decreases, given that the elasticity and the share of labor are constant, which is inversely related to the output markup μ_j.

This apparently direct positive relationship between the wage and the product price markup could be interpreted as if the larger the firm's rent, the larger the wage markup (as in Dobbelaere (2004)). However, according to our analysis, it might be just an underestimation of the true level of price–cost margins that is caused by the omission of direct effects of the wage bill on marginal costs (Bughin, 1993). As a matter of fact, finding a significant estimate for the wage markup parameter μ_{it}^{W} means that the workers' union has a degree of bargaining power, ϕ_{it}, which erodes the existing monopoly rents. Therefore, we expect price–cost margins and bargaining power parameters (both wage markup and bargaining elasticity) to be negatively related.

Finally, we include the labor elasticity, as expressed in (4.4), in the deflated revenue function (2.7). The resulting estimating equation assuming labor market frictions can be expressed as[7]

$$a. \qquad y_{it} = \gamma_{i0t} + \gamma_{iKt}k_{it} + \gamma_{iMt}m_{it} + (1-\mu_{it}^{W})s_{iLt}l_{it} - \frac{1}{\eta_j}q_t^j + \tilde{a}_{it} + \tilde{u}_{it}, \tag{2.13}$$

$$b. \qquad y_{it} = \gamma_{i0t} + \gamma_{iKt}k_{it} + \gamma_{iMt}m_{it} + \gamma_{iLt}l_{it} - \frac{1}{\eta_j}q_t^j + \tilde{a}_{it} + \tilde{u}_{it}, \tag{2.14}$$

$$c. \qquad q_{it} = \theta_{i0t} + \theta_{iKt}k_{it} + \theta_{iMt}m_{it} + \theta_{iLt}l_{it} + a_{it} + u_{it}^{y}. \tag{2.15}$$

Estimating equation b in (2.14) represents the deflated revenue function allowing for imperfect competition in the product market and perfect competition in the

[7] This functional form allows us to directly obtain an estimate of μ_{it}^{W} and c, along with the estimation of the other regression parameters. As a robustness check, we also specify the labor elasticity in the empirical application as $\gamma_{iLt} = s_{iLt} - \frac{\phi_{it}}{1-\phi_{it}}\frac{\Pi_{it}}{P_{it}(Q_{it})Q_{it}}$, and directly obtain an estimate of $\frac{\phi_{it}}{1-\phi_{it}}$ from the following estimating equation:

$$y_{it} - s_{iLt}l_{it} = \gamma_{i0t} + \gamma_{iKt}k_{it} + \gamma_{iMt}m_{it} - \frac{\phi_{it}}{1-\phi_{it}}\frac{\Pi_{it}}{P_{it}(Q_{it})Q_{it}}l_{it} - \frac{1}{\eta_j}q_t^j + \tilde{a}_{it} + \tilde{u}_{it}.$$

The bargaining parameter, ϕ_{it}, in the first specification, and the *wage markup*, μ_{it}^{W}, in the latter, are retrieved by exploiting the functional relation of equation (2.10), where:

$$\frac{W_{it} - \overline{W}_{it}}{W_{it}} = \hat{\mu}_{it}^{W} = \frac{\hat{\phi}_{it}}{1-\hat{\phi}_{it}}\frac{\Pi_{it}}{W_{it}L_{it}}$$

and standard errors are computed using the delta method. This yields similar empirical results.

labor market. Specification b could still represents the revenue function of a firm producing in a unionized labor market where the workers bargain only over the wage and let the employers determine the level of the employment, i.e., right-to-manage (Dobbelaere and Mairesse, 2011). Specification c in (2.15) represents the production function in the perfect competition case, and it used as benchmark.

Estimation of equation (2.13) can be done following several estimation approaches, under appropriate corresponding distributional assumptions, which will be discussed in the next section.

2.3 Estimation and identification strategies

In this section we briefly review the main problems concerning the estimation of a production function. Moreover, we discuss the advantages and drawbacks of some widely adopted estimation techniques, linking them to our structural framework of imperfect competition. The literature that we are going to investigate does not necessarily discuss our specification (2.13), but typically the version without imperfect competition in both output and labor market, i.e., with $\mu_{it}^W = 1$ and $1/\eta_j = 0$. The corresponding equation is given by

$$y_{it} = \gamma_{i0t} + \gamma_{iKt}k_{it} + \gamma_{iMt}m_{it} + \gamma_{iLt}l_{it} + a_{it} + u_{it}. \tag{2.16}$$

In empirical applications one usually assumes that the parameters are constant, i.e., $\gamma_{ikt} = \gamma_k$, for $k \in \{0, K, L, M\}$, while u_{it} is treated as a usual error term. The problem is the potential correlation between a_{it} and the inputs that are chosen at time t, in our case at least l_{it}. Assuming that the unobserved productivity is constant over time ($a_{it} = a_i$), the potential endogeneity between a_i and the inputs is controlled by exploiting the panel structure of the data, for instance, by using the fixed-effects estimator. However, if we believe that a_i evolves over time instead, using the fixed effects estimator does not solve the endogeneity problem, and alternatives have to be investigated.

One natural alternative is the use of an instrumental variables (IV) estimator, if instruments can be found that are correlated with the inputs, but uncorrelated with a_{it} and u_{it}. A natural candidate for such instruments are the input prices.

However, as discussed by Ackerberg et al. (2007), using input prices as instruments has not been uniformly successful in practice. Moreover, assuming market frictions in the input markets (as we do with respect to the labor market), might imply correlation between a_i and the input prices (in our case wages). Therefore, other instruments, or other estimation approaches have to be investigated.

2.3.1 Control function estimation approach

In their seminal paper, Olley and Pakes (1996) propose a control function approach to estimate production functions (which they considered without intermediate goods, and where labor is the only endogenous input). In particular, Olley and Pakes (1996) assume that the investment level is strictly monotonic in the scalar unobservable productivity level a_{it}. More precisely, they assume that investment i_{it} satisfies $i_{it} = j_t(a_{it}, k_{it})$, strictly increasing in a_{it}. It is then possible to invert the investment demand function. This yields a so-called control function, expressing productivity as a function of investment, along with other variables (in their case capital k_{it}): $a_{it} = j_t^{-1}(i_{it}, k_{it})$. By substituting out the unobserved productivity a_{it} using this control function, the resulting equation does not have endogeneity problems anymore.

However, when inverting the investment function, to guarantee the one-to-one mapping between firm-level productivity and the observable investment, the investment variable has to be strictly positive. As investments in the data are often zero, Levinsohn and Petrin (2003) also include intermediate goods as extra endogenous input in the production function. Then they propose the use of an intermediate input demand function, such as materials and energy demand functions, as a proxy for unobserved productivity: $m_{it} = f_t(a_{it}, k_{it})$. This yields an alternative control function, assuming that f_t is strictly increasing in m_{it}: $a_{it} = f_t^{-1}(m_{it}, k_{it})$.

In practice, the estimation approach suggested by Olley and Pakes (1996) and Levinsohn and Petrin (2003) is composed of two stages. In the first stage, one estimates the labor elasticity, along with the substituted replacement function, which can be approximated by a sufficiently high order polynomial in its arguments. However, due to the nonparametric character of the control function the capital and possibly material elasticities cannot be estimated in the first stage. Thus, in the first stage, focusing on the Levinsohn and Petrin (2003) approach, one runs a regression

$$y_{it} = \gamma_L l_{it} + \phi_t(m_{it}, k_{it}) + u_{it}, \tag{2.17}$$

with ϕ_t approximated by a higher order polynomial in m_{it} and k_{it}, where ϕ_t satisfies

$$\phi_t(m_{it}, k_{it}) = \gamma_K k_{it} + \gamma_M m_{it} + f_t^{-1}(m_{it}, k_{it}).$$

In the second stage, to identify the other production function parameters (capital and material), Olley and Pakes (1996) assume that the technological progress, a_{it}, depends on the information known by the firm i, at time $t-1$ (information set I_{it-1}). Under the additional assumption that a_{it} follows a first–order Markov process, the past realizations of a_{it} constitute the information set. In other words,

$$a_{it} = h(I_{it-1}) + \xi_{it} = h(a_{it-1}) + \xi_{it}, \tag{2.18}$$

where ξ_{it} represents the unanticipated innovative shock to productivity, assumed to be uncorrelated with material and capital in period $t-1$. Using the first stage estimates from (2.17) and assumption (2.18), we can write the production function (2.16) as:

$$y_{it} - \widehat{\gamma}_L l_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + h[\widehat{\phi}_{it-1} - \gamma_0 - \gamma_K k_{it-1} - \gamma_M m_{it-1}] + \xi_{it} + u_{it}, \tag{2.19}$$

where

$$\widehat{\phi}_{it-1} = \widehat{\phi}_t(m_{it-1}, k_{it-1}). \tag{2.20}$$

Approximating $h(\cdot)$ with a flexible polynomial, (2.19) can be estimated, using lagged materials m_{it-1} as instrument for materials m_{it}, since m_{it} might be correlated with ξ_{it}.

A restricting assumption of the Olley and Pakes (1996) model concerns the timing and dynamic nature of inputs, i.e., some inputs are more "dynamic in nature" than others. In particular, when selecting the state variables that enter the firm's expected discounted profit maximization, Olley and Pakes (1996) define capital as a dynamic input, i.e., its choice of the current period affects the choice of the next period. Labor (and in case of Levinsohn and Petrin (2003) also materials), on the other hand, is assumed to be a more flexible and non–dynamic input, therefore, it is implicitly assumed that there are no labor (and material) adjustment costs.

2.3.2 Ackerberg et al. (2006) critique and solutions

According to Ackerberg et al. (2006), an important drawback of the Olley and Pakes (1996) approach in the context of (2.16) arises form collinearity between labor and the polynomial in material (or investment) and capital. As a matter of fact, in the first stage, the labor coefficient could be unidentified, if the regressor l_{it} does not have any sample variability that is independent of the other regressors. To see this, let us consider the approach of Levinsohn and Petrin (2003) which assumes that m_{it} and l_{it} are perfectly flexible inputs, chosen simultaneously. In the context of equation (2.16), i.e., under perfect competition and with k_{it} an inflexible input factor, the firm's state variables at time t are given by a_{it}, k_{it}, P_t, W_t, and Z_t, where the output price P_t, wage W_t, an the material price Z_t are assumed to be constant over firms (assuming perfect competition in the three markets). The demand of material and labor are then functions of these state variables, i.e., $m_{it} = f(a_{it}, k_{it}, P_t, W_t, Z_t) \equiv f_t(a_{it}, k_{it})$, and $l_{it} = q(a_{it}, k_{it}, P_t, W_t, Z_t) \equiv q_t(a_{it}, k_{it})$. Therefore, they both depend on the same state variables, a_{it} and k_{it}, and $l_{it} = q_t(f_t^{-1}(m_{it}, k_{it}), k_{it}) = s_t(m_{it}, k_{it})$, leaving the labor input as a time-varying function of material and capital. But then there is no independent firm-level source of variation that could help identify the labor elasticity coefficient in the first stage. In addition, Ackerberg et al. (2006) show that these problems remain intact, even if one is able to derive f_t^{-1} explicitly, assuming a Cobb-Douglas production function: again, one cannot the labor elasticity coefficient in the first stage.

Ackerberg et al. (2006) discuss two possible ways to break this collinearity issue. One way is to assume that the firm makes optimization errors only in the choice of labor (optimization error in material adds an additional unobservable, violating the scalar unobservability assumption, needed to construct a control function). This error does not enter the production function and will move l_{it} around independently of the control function.

The second option is to make specific timing assumptions that would justify the independency of cross–sectional variation of l_{it}, conditional on the choices of materials and capital. In particular, Ackerberg et al. (2006) propose to assume that, next to capital, also labor is a state variable, i.e., not fully flexible, and that labor is chosen somewhere after the choice of capital (at time $t-1$) and before the choice of materials (at time t). Given this timing assumption, Ackerberg et al. (2006) proceed with a two–step estimation approach, where the labor coefficient is identified only in the second stage, together with the capital coefficients. As estimation approach, Ackerberg et al. (2006) propose to use two equations, namely, first

$$y_{it} = \gamma_K k_{it} + \gamma_M m_{it} + \gamma_L l_{it} + g(l_{it}, m_{it}, k_{it}) + u_{it}, \tag{2.21}$$

with g a polynomial in terms of l_{it}, m_{it}, and k_{it}, and assuming that the expectation of u_{it} conditional on time t information is equal to zero. The second equation follows from (2.18), and is given by

$$g(l_{it}, m_{it}, k_{it}) = h[g(l_{it-1}, m_{it-1}, k_{it-1})] + \xi_{it}, \tag{2.22}$$

with h a univariate polynomial, and assuming that the expectation of ξ_{it} conditional on time $t-1$ information is equal to zero. This approach is based on Wooldridge (2009) in order to obtain efficient GMM estimates and standard errors in one step. In fact, Wooldridge (2009) uses the second equation substituted in the first one, yielding

$$y_{it} = \gamma_K k_{it} + \gamma_M m_{it} + \gamma_L l_{it} + h[g(l_{it-1}, m_{it-1}, k_{it-1})] + \xi_{it} + u_{it}, \tag{2.23}$$

together with the assumption that the expectation of $\xi_{it} + u_{it}$ conditional on time $t-1$ information is equal to zero.

2.3.3 Bond and Söderbom (2005) critique and IV solution

Bond and Söderbom (2005) also illustrate the identification issues concerning the parameters of a Cobb-Douglas production function, when these are assumed to be perfectly flexible. Indeed, when the input prices are common to all firms and inputs are chosen optimally with no adjustment frictions, assuming perfect competition on output and input markets, the levels of the inputs are perfectly collinear (in the sense of linear regression) with each other and the productivity. Bond and Söderbom (2005) show how this collinearity problem is not solved by assuming that one of the inputs, i.e., the level of capital, is predetermined (chosen in the previous period). According to these authors, the only assumption that guarantees the identification of the structural parameters is a positive and exogenous variation in input prices (such as one deriving from adjustment costs). However, according to Bond and Söderbom (2005), the presence of unobserved variation across firms in prices might rule out the control function approach, which is assumed to be common to all firms.

To break the collinearity issue, Bond and Söderbom (2005) consider the dynamic problem of a firm maximizing the current and future profits in the presence of firm-specific adjustment costs and productivity shocks. Simulating data for a two-factor Cobb-Douglas production function, Bond and Söderbom (2005) show that the production function coefficients are identified provided that there are adjustment costs for each input and that these inputs are subject to different levels of adjustment costs. Given the empirical evidence of the presence of such adjustment costs presence, Bond and Söderbom (2005) suggest the use of instrumental variables methods to consistently estimate the parameters of the production function. In particular, the lagged level of inputs constitute informative instruments, as the presence of cost frictions makes the variation of these inputs persistent.

2.3.4 Gandhi et al. (2011) critique and solution

As later stressed by Gandhi et al. (2011), if one is not willing to assume the presence of adjustment frictions in all inputs, the nonparametric identification of the production function fails in the presence of flexible inputs. In particular, Gandhi et al. (2011) show how, in the perfect competition case, when the material is assumed to be a flexible input, its elasticity suffers from an identification problem deriving from the fact that there is no source of cross-sectional variation in m_{it} independent of the firm's remaining productive inputs (l_{it}, k_{it}, a_{it}). Moreover, according to Gandhi et al. (2011) the collinearity problem described by Ackerberg et al. (2006) is not solved. To argue this, they consider equation (2.23) of the Wooldridge (2009) approach, together with the (specialized) moment condition

$$E(\xi_{it} + u_{it} | l_{it-1}, m_{it-1}, k_{it-1}) = 0.$$

The lagged value m_{it-1} is intended to act as an instrument for the endogenous m_{it}. However, m_{it-1} alone cannot act as an excluded variable and be a valid instrument for m_{it}, since m_{it-1} is already included in (2.23) as an additional regressor, and, moreover, it is fully collinear with l_{it-1} and k_{it-1}.

Gandhi et al. (2011) propose a solution that identifies the coefficient of a flexible input and is based on a transformation of the firm's first order condition for flexible inputs, i.e., material. The key idea behind their approach is to combine the transformation of the FOC for material with the idea of measurement error in output (unanticipated productivity shock). In particular, they obtain a simple "share regression" model, $\log(s_{iMt}) = \log(\theta_M) - u^q_{it}$, and perform non-parametric regression of $\log(s_{iMt})$ on all inputs levels (since $\log(\theta_M)$ depends on all productive inputs). Since the ex-post shock, u^q_{it}, is assumed to be independent of the three inputs, the non-parametric share regression identifies both material elasticity and u^q_{it}.

2.3.5 Estimation using market imperfections

In this subsection we describe and motivate our three estimation approaches, also referring to the discussion in the previous subsections. Our starting point is equation (2.13), assuming that the structural parameters are drawn from probability distributions that do not vary over firms or time, i.e., the resulting estimation equation is

$$y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + (1 - \mu^W) s_{iLt} l_{it} - \frac{1}{\eta_j} q^j_t + \tilde{a}_{it} + \tilde{u}_{it}. \qquad (2.24)$$

In our specification, at least, l_{it} is fully flexible. According to Bond and Söderbom (2005) this might result in perfect collinearity in a Cobb-Douglas production framework under perfect competition in output and input markets. However, due to the imposed market imperfections, we do not have this collinearity problem. First, as independent variable we do not have l_{it} in equation (2.24), but $s_{iLt}l_{it}$. Second, l_{it} has to satisfy (4.21)–(4.22), avoiding that l_{it} depends in a linear way on $\widetilde{a}_{it}$ and the other inputs, as derived by Bond and Söderbom (2005) in the perfect competition case. Moreover, if k_{it}, m_{it}, or both are fully flexible as well, then (4.21)–(4.22) will again avoid perfect collinearity.

Next, we describe our estimation approaches, in combination with the corresponding additional distributional assumptions. As first estimation approach, we shall use the fixed effects (or within) estimator, which requires as additional assumption that $\tilde{a}_{it}$ is not time dependent, i.e., $\tilde{a}_{it} = \tilde{a}_t$.

As second estimator, we shall use Wooldridge (2009)'s estimator related to the Levinsohn and Petrin (2003) control function approach. In this case we add (2.18), together with the assumption that ξ_{it} has mean zero, conditional on all time $t-1$ information. To be able to apply the control function approach, we "derive" $\widetilde{a}_{it}$ as a (nonlinear) function of the inputs l_{it}, k_{it}, and m_{it}, and the state variables, using our set-up. When doing this, the collinearity problems described by Ackerberg et al. (2006) for the Cobb-Douglas production function do not show up, due to the market imperfections. Moreover, the assumption that the production function is a Cobb-Douglas production function avoids the *nonparametric* identification problems raised by Gandhi et al. (2011). However, we have to assume that $\widetilde{a}_{it}$ is only a (nonlinear) function of the inputs l_{it}, k_{it}, and m_{it}, constant over firms an time. We then use (2.23), updated to our specification, i.e.,

$$y_{it} = \gamma_K k_{it} + \gamma_M m_{it} + (1 - \mu^W) s_{iLt} l_{it} - \frac{1}{\eta_j} q_t^j + h[g(l_{it-1}, m_{it-1}, k_{it-1})] + \xi_{it} + \tilde{\tilde{u}}_{it}, \tag{2.25}$$

together with the distributional assumption that $\xi_{it} + \tilde{\tilde{u}}_{it}$ has mean zero, conditional on all information at time $t-1$. We estimate (2.25), with g and h specified as third order polynomials, and using as instruments the lagged inputs l_{it-1}, k_{it-1}, m_{it-1}, and their higher order and interaction terms, up to the third order. This choice of instruments in particular makes sense if there is some persistence over time in the inputs, for instance, via k_{it}. Given persistence in k_{it}, there will also be persistence in the other inputs, due to the (nonlinear) dependence between the inputs.

As third estimation approach, we shall use the IV estimator. Here, we use (2.24), and make the assumption that $\tilde{a}_{it} + \tilde{\tilde{u}}_{it}$ has mean zero, conditional on all information available at time $t-1$. We estimate (2.24), using GMM, with as instruments the first and second order lags of l_{it}, k_{it}, and m_{it}. Again, given persistence in, for instance, k_{it}, there will also be persistence in the other inputs, making that the lagged inputs are correlated with the current inputs. However, although the IV approach has the advantage that it does not require the assumption of a control function, constant over firms and time, the cost of adopting this estimation approach is that one does not allow for the possibility that the unobserved productivity could be correlated with past choices of inputs.

2.4 Data

We extract data from Statistics Netherlands for the years 1989-2008. As an output measure, we use the deflated value of gross output Y_{it} ($\equiv \frac{Q_{it}P_{it}}{P_t^j}$) of each firm i in sector j in period t. Labor (L_{it}) refers to the number of employees in each firm for each year,[8] collected in September of that year. The corresponding wages W_{it} include gross wages plus salaries and social contributions before taxes. The costs of intermediate inputs ($Z_{it}M_{it}$) include costs of energy, intermediate materials, and services. The unit user costs R_{it} (of capital stock K_{it}) are calculated as the sum of the depreciation of fixed assets and the interest charges.

Table 2.1: Descriptive Statistics

variable	mean	sd	median	p25	p75	N
Y_{it}	26531.561	1.36e+05	6360.172	2895.117	16657.439	60672
L_{it}	112.162	303.621	50	28	105	”
M_{it}	18100.772	98131.263	3885.728	1572.234	10905.304	”
K_{it}	1671.484	16171.453	247.561	95.557	722.472	”
s_{iLt}	0.270	0.127	0.179	0.258	0.345	”
s_{iMt}	0.615	0.147	0.520	0.620	0.716	”
s_{iKt}	0.045	0.044	0.019	0.035	0.058	”
Q_t^j	0.940	0.153	0.853	0.966	1.039	”

Note: p25 and p75 are, respectively, the 25th and the 75th percentile.

[8]For each enterprise, jobs are added and adjusted for part-time and duration factors, resulting in number of men/years expressed as Full Time Equivalents (FTEs)(*Source:* Statistics Netherlands)

The nominal gross output and intermediate inputs are deflated with the appropriate price indices from the input-output tables available at the NACE rev. 1 two-digits sector classification.[9] For capital, we use a two-digit NACE deflator of fixed tangible assets calculated by Statistics Netherlands.

The data extracted from the Production Survey (PS) constitutes an unbalanced panel data of 6727 firms (with a minimum of 2001 firms in 2001 and a maximum of 5607 enterprises in 2006 and 1997) with 65866 observations spanning over 20 years and over 21 industries. We exclude from the sample firms producing for less than two consecutive years.[10] Also, firms with missing data on one of the variables used in the empirical analysis are omitted. We exclude firms exhibiting inputs growth of more than 200 percent or less than -50 percent (3822 observations dropped). We also exclude firms with an output growth of more than 300 percent or less than -90 percent (1372 observations). The resulting sample consists of 60672 observations (6718 firms).

Throughout our sample period, the PS surveys included some changes in their population designs resulting in an unbalanced panel. As a result, we cannot distinguish whether the entry or exit rates of firms resulted from survey response behavior or real economic structural behavior. The number of firms (N) for each NACE rev. 1 industry is calculated by Statistics Netherlands. Table 2.9 reports the sectors that were chosen with a corresponding NACE two-digit code and the corresponding number of firms.

Table 2.1 reports the means, medians, standard deviations, and first and third quartiles of the included data for our main variables. In particular, a summary of the of deflated revenues and of the inputs (in thousands Euros), along with input shares in revenue, is presented. The input shares are constructed by dividing respectively the (undeflated) firm input cost by the firm undeflated revenues. As

[9]NACE Rev. 1 is a 2-digit activity classification which was drawn up in 1989. It is a revision of the General Industrial Classification of Economic Activities within the European Communities, known by the acronym NACE and originally published by Eurostat in 1970.

[10]The numbers of firms for each number of observation are: 700(2), 1806(3), 2556(4), 3400(5), 2760(6), 3073(7), 2256(8), 3222(9), 2730(10), 3168(11), 3396(12), 4134(13), 4032(14), 3810(15), 3856(16), 3995(17), 4806(18), 3515(19), 3520(20), where the number of observation per firm is reported between brackets.

one can see, the dispersion of the logarithms of deflated output and inputs is considerably large.[11]

During 1989-2008, the capital input constitutes 4.5 percent of gross output on average. The mean share of labor is 27 percent, and intermediate inputs constitute more than half of gross output (61.5 percent). Moreover, the relative dispersion of all these variables is considerably large, especially for the share of capital. The exit rate, non reported here, is quite small (2.7 percent) and 75 percent of the firms have been active on the market for 3 to 10 years.

2.5 Empirical results for the complete sample of Dutch firms

In this section we present results for the entire manufacturing industry over the period 1989-2008, without looking at the potential heterogeneity in the structural parameters across firms and/or through time, using the random coefficients framework. Section 5.1 explores the empirical results where the relevant parameters are allowed to differ among product segments.

Table 2.2 reports the estimated parameters of interest of the production function for the whole manufacturing industry. The table is organized per estimation approach: fixed effects (Within), instrumental variables (IV), Wooldridge (2009)'s estimations of the Levinsohn and Petrin (2003) model (Wool.–LP). As already reported, we estimate the latter model parametrically, proxying the unknown functions $g_t(\cdot)$ and $h(\cdot)$ in equation (2.25) with polynomials of third order. The independent variable $s_{iLt}l_{it}$ is calculated by taking for the labor production share s_{iLt} the sector specific labor production share s^j_{Lt}, with $j = j_i$. For each econometric approach, we consider three model specifications: a (the production function featuring both labor and product market imperfect competition), b (product market imperfect competition only), and c (no corrections for omitted price bias and labor imperfect competition). The model specifications b and c are included for comparison purposes only.

[11] Averages over time and standard deviations for each sector are reported in Table 2.9 in Appendix C.

With every estimation technique, the results confirm the hypothesis of simultaneous output and labor markets imperfections. Columns a report evidence of markups larger than one, rent–sharing parameters and wage markups larger than zero. In line with Bughin (1993, 1996), Dobbelaere (2004), Abraham et al. (2009), and Dobbelaere and Mairesse (2011)), we find that excluding the rent sharing parameter (columns b) leads to an underestimation of the product markup. The omitted output price bias (arising from the use of deflated gross output instead of output in volumes) is evident when comparing columns b and c. All inputs elasticities are biased downwards when we do not implement the correction for the omitted output price bias (as suggested by Klette and Griliches (1996) and De Loecker (2011)). Indeed, given that inputs and output are positively correlated and output and price are negatively correlated, we expect the correlation between inputs and firm–level price differences to be negative. These downward biases are significant for the labor and material coefficients, but not statistically significant for the capital coefficient estimated using the control function approach (Wool.–LP). De Loecker (2011) also finds similar results. In general, the variation of the estimates of the capital coefficient when introducing the demand shifter is much smaller compared with the other input elasticities. Additionally, omitting the price variable leads to the underestimation of the scale elasticity. Our estimates of the production function are not always in line with those found in this literature. De Loecker (2011), using data from the Belgian textile industry, corrects for both simultaneity (using the Olley and Pakes (1996)'s procedure) and omitted price bias estimates and finds the estimated output markup equal to 1.45 against our 1.09. Indeed, comparing his results with our column b of the Wooldridge's one step efficient estimation of Levinsohn and Petrin (2003), we find much smaller production function coefficients (also smaller standard errors) and output markup. De Loecker (2011) finds the inputs elasticities of labor, material, and capital equal to 0.307, 0.906, 0.150, respectively.[12]

Dobbelaere and Mairesse (2011), with a first–difference OLS estimation of a production function featuring an efficient bargaining in the labor market, find decreasing returns to scale (0.792) versus our slightly increasing scale elasticity (1.031, column a of the within estimator). On the other hand the labor and

[12]The standard errors of nonlinear combination of the estimated coefficient are computed using the Delta method.

capital elasticities are of approximately the same small magnitude. Their output markup is equal to 1.102, while with our data on the Dutch manufacturing industry, we find this to be larger (1.201). The bargaining parameter found by Dobbelaere and Mairesse (2011) with the OLS estimator is much larger (0.552) than what Dobbelaere (2004) and we find, with the Arellano and Bond (1991) and IV estimators, respectively. Indeed, the bargaining elasticity is equal to 0.482 (within estimator), while with the IV estimator we find a bargaining power of 0.244, which coincides with what Dobbelaere (2004) finds. Dobbelaere and Mairesse (2011) also perform a Blundell and Bond (2000)'s estimation of the production function. Under this estimation approach, Dobbelaere and Mairesse (2011) still find a low estimate of capital elasticity and slightly decreasing returns to scale (0.033 and 0.969, respectively). Column a of the IV estimation reports capital and scale elasticities of 0.080 and 1.301, respectively. On the other hand, the output markup and bargaining parameter are much larger than what we find. The markup is equal to 1.383 against our 1.284; the estimated bargaining parameter is 0.552, while we find this to be equal to 0.244. However, the results of Dobbelaere (2004) for Belgian data are somewhat inconsistent with the results of Dobbelaere and Mairesse (2011), who use French data. The union's bargaining power parameter, according to McDonald and Suen (1992) framework, is positively dependent on the ratio between the trade union density[13] and the unemployment rate. As this ratio is, at the aggregate level, higher for Belgium than for France (roughly, 6.2 and 0.9, respectively; source: OECD.Stat), one would expect to find higher bargaining power in Belgium rather than in France. On the other hand, the Netherlands report a trade union density/unemployment rate ratio of approximately 4. We would then expect to find a somewhat smaller bargaining coefficient, rather than the exact same magnitude (if not larger, as we move to the control function approach estimates). We find consistent results with the (weighted) average of the workers' power parameters over sectors. Indeed, except for the within estimator, we find bargaining elasticities ranging between 0.196 and 0.217.

Along with the bargaining parameter, we are able to provide, within the production function framework, an estimate of the wage markup. We notice how

[13]Trade union density corresponds to the ratio of wage and salary earners that are trade union members, divided by the total number of wage and salary earners (OECD Labour Force Statistics).

the wage markup estimate is upward biased, when using the within estimator (63.6%). When adopting a GMM framework, the wage markup fluctuates between 21.9% (IV) and 23.7% (one step Levinsohn and Petrin (2003)). These results are comparable with Aidt and Tzannatos (2008)'s review of wage markups in high–income economies, which are between 0 and 25%.

In addition to the parameters mentioned above, we also report the profit ratio parameter, which can be expressed as the estimated product markup divided by the estimated elasticity of scale, $\frac{\hat{\mu}}{\hat{\theta}}$. A profit ratio parameter larger than one could be due to either imperfect competition or to decreasing returns to scale.

2.6 Across-Industry Estimates

The nature of the underestimation discussed in section 2.3 could be structural or computational. If structural, i.e., the workers tend to gain higher wage rents in those sectors where there is less competition, we expect to find a positive association between the wage markup μ^{Wj} and the output markup μ_j (see equation 4.4). On the other hand, if the existing monopoly rents are eroded by some bargaining power of workers' union, the bias derives from an underestimation of the true level of price–cost margins (Bughin, 1993, see) and we expect a negative correlation between the two parameters.

Given the evidence of sectoral specificity of capital and labor[14] (Dosi, 1999; Ramey and Shapiro, 2001), we investigate the heterogeneity of the manufacturing industry, by studying across-industry firms' production behavior.

For each of the 21 industries, we estimate equation (2.13) with and without the extension to labor imperfections (model specifications a and b, respectively). Year dummies are always included. Moreover, the production index is constructed as in De Loecker (2011), by proxying the total demand for a 6-digit sector $\tilde{j}$ with a (market share) weighted average of deflated revenue, $q_t^{\tilde{j}} = \sum_i^{N_{\tilde{j}}} ms_{it}y_{it}$. Table

[14] The "social embeddedness" of firms' routines and strategies is likely to be driven by socially specific factors, such as the nature of the local labor markets, workforce training institutions, financial institutions. Furthermore, Ramey and Shapiro (2001) suggest significant sectoral specificity of physical capital and substantial costs of redeploying the capital.

2.4 reports the within estimates of the relevant parameters, namely the output and wage markups, $\widehat{\mu}_j$ and $\widehat{\mu}^{Wj}$, respectively, and the bargaining elasticities, $\widehat{\phi}^j$. Table 2.5 displays the IV estimated coefficients, while Tables 2.6 reports the structural estimates obtained using the control function approach.

Testing the hypothesis of heterogeneity across sectors yields the conclusion that all the structural parameters statistically differ from sector to sector, and are sensitive to the estimation technique. This confirms the assumption of sectoral specificities. Each sector has its own functioning, and the firm belonging to a specific sector adopts a different production strategy compared to a firm in another sector.

Quite consistently with what we found for the whole manufacturing industry, excluding imperfections on the labor market leads to an underestimation of the markups for the majority of sectors. When assuming imperfect competition on the labor market and on the output market, the output market markups range from 0.940 (sector 27, metals) to 2.730 (sector 25, rubber and plastic products) for the within estimator; from 0.899 (sector 19, textile and leather products) to 3.380 (sector 20, wood) for the IV estimator; from 0.869 (sector 19) to 3.446 in the manufacturing sector of wood (Wool.–LP estimator). On the other hand, when disregarding the possibility to have frictions on the labor market, the output price markups range from 0.857 (sector 27) to 1.439 (sector 25) for the within estimator; from 0.921 (sector 19) to 1.632 (sector 23, coke, petroleum and nuclear fuel) for the IV estimator; from 0.853 (sector 32) to 1.288 (sector 18) for the Wool.–LP estimator.

With every estimation technique, we find negative correlation between the parameters μ_j and the labor market friction parameters μ^{Wj} and ϕ^j. The Spearman's rank correlation coefficient becomes statistically significant as we perform an IV, or Wool.–LP. This confirms the hypothesis of a pure computational bias of the true level of the output markup μ_j, possibly caused by the misspecification of the marginal costs, as we are omitting the direct effects of wage rigidities. Therefore, firms share their monopoly rents with labor unions. This result is in contrast with the findings of Crépon et al. (2002), Dobbelaere (2004), and Dobbelaere and Mairesse (2011) who find positive correlation. Dobbelaere (2004) interprets this positive correlation between labor bargaining and output market power as the effect of the exit of firms. In particular, she guesses that strong unions reduce the firms' share of rents, forcing some of the firms to exit the market, therefore decreasing the degree of market power. Another explanation she provides deals with the fact that stronger unions are attracted by those sectors where rents can be extracted.

Both these two interpretations have problems. The first interpretation builds on the premise of a static setting, which does not allow for the dynamic aspects of competition (such as the implications of selection bias and reallocation effects). The second interpretation concerns more the profitability of the firm, rather than its level of price–cost margin. A more profitable firm can attract workers that are able to extract some of the surplus. But a higher markup does not necessarily mean that the enterprise is profitable, as it does not take into account relative cost efficiencies (see Boone and van der Wiel (2007); Boone (2008); Griffith et al. (2008) for a discussion on relative profits and relative cost efficiencies). Therefore, we tested the correlation between the wage markup and the relative profits measure (computed as in Boone and van der Wiel (2007)). We find that indeed these two measures, profit elasticity and union power, are positively correlated ($\rho = 0.46$, significant at the 5% level. Results of the profit elasticities per sector are not reported, but available upon request.)

2.7 Impact on TFPG

In this paper we propose a measure of TFP growth (TFPG) derived from estimating a production function which accounts for both imperfect competition on the labor market and on the output market as derived in Section 2. The TFP measure is computed as

$$TFP_{it} = \hat{a}_{it} \equiv \hat{\mu_j} \left[y_{it} - (\hat{\gamma}_K k_{it} + \hat{\gamma}_M m_{it} + (1 - \hat{\mu}^{Wj}) s_{iLt} l_{it} - \frac{1}{\hat{\eta}_j} q_t^j) \right]. \quad (2.26)$$

To compute the TFP growth index, we follow De Loecker and Konings (2006) and take a employment based share weighted firm-level TFPG, where the shares are simply $\pi_{it}^j = L_{it} / \sum_i L_{it}^j$.

We also derive a measure of TFPG following equation ((2.26)), allowing for the structural parameters to vary across sectors. Table 2.3 reports the weighted average TFPG percentage rates for every estimation approach and for all three specifications (a, b, and c). We find that correcting for the omitted output prices, therefore taking into account the possibility for the firms to set prices above their marginal costs, leads to decreases of the TFPG percentage rates of 2.3% (within), 4.5% (IV), and 1.2% (Wool.-LP). Independently from the estimation technique used, when assuming both bargaining on the labor market and imperfect competition on the output market, we find larger TFPG rates (column a). In particular, with the within estimator we find a 47.6% higher TFPG rate than the TFPG rate estimated under the sole assumption of imperfect competition on the product market. With the IV, and the Wool.–LP, we find 158.2% and 122.8% higher TFPG rates, respectively. Moreover, the size of the firm, measured as number of employees, seems to be negatively correlated with the TFPG rates when using the IV or control function approaches. In particular, we find that firms with less than 50 employees are the most productive ones.

Figure 2.1 reports the three time series of the TFPG rates under the three different market structure assumptions (imperfect competition on both output and labor markets, a, imperfect competition on the output market, b, and perfect competition, c). Each panel reports the TFPG index under the three different estimation strategies. The first striking feature of the estimates of the firms' average productivity growth when assuming imperfect competition on both markets is its larger variance, with growth rates ranging from -3.96% to 9.95% (IV estimation approach). Indeed, when considering the model specification b or c, we find ranges of values between -1.09% and 3.07%, which are in line with the TFPG rated reported by the OECD[15] for the Dutch manufacturing sector. Moreover, with each estimation approach we find evidence of a positive time trend only when we consider both labor and product market imperfections.

Table 2.7 reports the percentage growth rates estimated using IV and the Wool.-LP approach for the functional assumptions a and b. We confront these two estimation approaches as we want to investigate the consequences of allowing the productivity to be distributed as an AR(1) process. As one can see, when assuming wage premia and price markups (model a), omitting or including the

[15] `http://www.oecd.org/statistics/productivity`.

dynamics of the productivity process (IV or Wool.-LP) does not make a difference in the TFPG rates, on average. On the other hand, when assuming only product imperfect competition, the Levinsohn and Petrin (2003) estimator yields much lower growth rates. Looking at the first two columns of the table, we see that the sectors that display very high TFPG rates (around 3%) wi are textile and leather products, metal products, machinery and equipment, and transport equipment. Sectors that have a TFPG rate between 2% and 3% are rubber and plastic products, electrical and optical equipment, and transport equipment.[16]

Table 2.8 displays the Spearman's correlation coefficients between TFPG rates and the structural parameters describing the imperfect competition on the labor and on the output market. When considering the possibility of a first-order Markov productivity process, the weighted average productivity growth rate seems to be positively associated with both labor market frictions parameters, and with the output price mark-ups; while with the IV estimation of the TFPG, we find a negative correlation between productivity growth, output, and labor input price mark-ups.

Figure 2.2 plots the TFPG rates obtained from the two estimation approaches (IV and Wool.-LP) against the workers bargaining parameter ϕ^j for each sector. It is easy to detect the positive correlation between the productivity growth and the bargaining parameter in the second plot. The economic benefits of unions could be found in the worker–manager cooperation. Indeed, unions can increase firms' productivity by "shocking" the management into better production practices (Aidt and Tzannatos, 2008). Masayuki and Morikawa (2010) empirically analyze the relationship between labor union and firm performance by using data on 4000 Japanese firms in both the manufacturing and non-manufacturing sector. The presence of labor unions has statistically and economically significant positive effects on firm productivity. Indeed unions may enhance productivity when these contribute to the strengthening of the firm's competitiveness, by fostering cooperation between labor and management.

[16] For the corresponding NACE codes of the sectors, see Table 2.9.

2.8 Conclusion

This paper explores some of the econometric issues regarding the empirical analysis of a Cobb-Douglas production function in an imperfect competitive setting. We consider to what extent the estimated unobserved firm–level productivity growth (TFPG) is sensitive to different model assumptions and to different econometric approaches.

Using firm–level data on 21 Dutch manufacturing sectors for the period 1989-2005, we investigate the possibility of having a manufacturing industry featuring both labor and output market distortions. In particular, we address the potential endogeneity issues concerning the simultaneity, the omitted output price, and the collinearity among input factors of production. We find that, omitting the evidence of workers' wage bargaining power, leads to a significant underestimation of the product markup (up to 16%). Omitting the potential correlation between input choices and firm-level price deviations yields downward biased input elasticities.

Moreover, we review different estimation techniques, and we propose an identification strategy that relies on the presence of imperfect competition in the flexible input market, namely, the labor input market. In particular, we compare three different estimation approaches that identify the structural parameters of the revenue production function, namely, the fixed-effect estimator, the control function approach, and the IV approach. With every estimation technique, our results confirm the hypothesis of simultaneous output and labor markets imperfections. Output price markups are estimated between 1.201 (within) and 1.284 (IV). When using the within estimator, both the wage markup and the bargaining elasticity estimates are upward biased (63.6% and 48.2%, respectively); while when adopting a GMM framework, the wage markup fluctuates between 21.9% (IV) and 23.7% (one step Levinsohn and Petrin (2003)).

When testing the hypothesis of heterogeneity across sectors, we find, consistently with what we found for the whole manufacturing industry, that excluding imperfections on the labor market leads to an underestimation of the markups for the majority of sectors. Moreover, we find negative correlation between the output price markup parameters and the labor market friction parameters, across the sectors. This confirms our hypothesis of a pure computational bias of the true level of the output markup, possibly caused by the misspecification of the marginal costs, as we are omitting the direct effects of wage rigidities. Therefore, firms share their monopoly rents with labor unions.

Additionally, we consider to what extent the estimated unobserved productivity is sensitive to the different model specifications and to the different approaches to identify the structural parameters of the model. We find that, when assuming wage and price markups, considering or not the dynamics of the productivity process (IV or Wool.-LP) does not make a difference in the TFPG rates, on average. On the other hand, when assuming only product imperfect competition, the Levinsohn and Petrin (2003) estimator yields much lower growth rates. With the Levinsohn and Petrin (2003) estimator the weighted average productivity growth rate seems to be positively associated with both labor market frictions parameters, and with the output price mark-ups; while with the IV estimation of the TFPG, we find a negative correlation between productivity growth, output and labor input price mark-ups. Both estimators, however, yield a significant and positive relation between productivity growth and the bargaining parameter. Therefore, the presence of labor unions has statistically and economically significantly positive effects on firm productivity. Indeed unions may enhance productivity when these contribute to the strengthening of the firm's competitiveness, by fostering cooperation between labor and management.

2.9 Appendix

Tables and Figures

Table 2.2: Results for the whole manufacturing industry

	Within			IV			Wool.–LP		
	a	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*
$\hat{\theta}_L$	0.118	0.240	0.216	0.271	0.237	0.213	0.258	0.225	0.201
	0.002	*0.003*	*0.002*	*0.011*	*0.011*	*0.006*	*0.010*	*0.007*	*0.006*
$\hat{\theta}_M$	0.864	0.760	0.680	0.951	0.811	0.732	0.912	0.781	0.716
	0.006	*0.005*	*0.001*	*0.024*	*0.012*	*0.004*	*0.021*	*0.011*	*0.004*
$\hat{\theta}_K$	0.048	0.036	0.032	0.080	0.068	0.061	0.062	0.056	0.051
	0.001	*0.001*	*0.001*	*0.004*	*0.003*	*0.003*	*0.004*	*0.003*	*0.003*
$\hat{\theta}$	1.031	1.036	0.929	1.301	1.116	1.006	1.233	1.061	0.973
	0.008	*0.007*	*0.002*	*0.034*	*0.016*	*0.002*	*0.030*	*0.021*	*0.005*
$\hat{\mu}$	1.201	1.120	–	1.284	1.109	–	1.258	1.091	–
	0.008	*0.007*		*0.033*	*0.016*		*0.029*	*0.014*	
$\frac{\hat{\mu}}{\hat{\theta}}$	1.165	1.081	–	0.986	0.994	–	1.020	1.028	–
$\hat{\mu}^W$	0.636	–	–	0.219	–	–	0.237	–	–
	0.006			*0.022*			*0.020*		
$\hat{\phi}$	0.482	–	–	0.244	–	–	0.191	–	–
	0.001			*0.001*			*0.024*		

Note: Sample period 1989-2008; dependent variable: log. gross deflated output y_{it}.
The estimated structural parameters are retrieved as the following: $\hat{\theta}_L$: sample mean of $\hat{\mu}(1-\hat{\mu}^W)s^j_{Lt}$, $\hat{\theta}_M = \hat{\mu}\hat{\gamma}_M$, $\hat{\theta}_K = \hat{\mu}\hat{\gamma}_K$, $\hat{\theta} = \hat{\theta}_L + \hat{\theta}_M + \hat{\theta}_K$, $\hat{\mu} = 1/(1+\frac{1}{\hat{\eta}})$, $\hat{\phi}$: sample mean of $\frac{1}{(1+\Pi_{it}/W_{it}L_{it}\hat{\mu}^W)}$.
a: $y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + (1-\mu^W)s^j_{Lt} l_{it} - \frac{1}{\eta} q^j_t + \tilde{a}_{it} + \tilde{\tilde{u}}_{it}$.
b: $y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + \gamma_L l_{it} - \frac{1}{\eta} q^j_t + \tilde{a}_{it} + \tilde{\tilde{u}}_{it}$.
c: $y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + \gamma_L l_{it} + a_{it} + u_{it}$.

Table 2.3: Weighted average TFP growth rates for the entire manufacturing sector (1989–2008)

		a	b	c
$Within$	$L \leq 50$	1.71	1.48	1.37
	$L \leq 250$	1.34	0.83	0.90
	$L > 250$	1.52	1.13	1.11
	avg	1.27	0.86	0.88
IV	$L \leq 50$	2.13	1.08	1.00
	$L \leq 250$	1.41	0.35	0.46
	$L > 250$	1.39	0.30	0.40
	avg	1.73	0.67	0.70
$Wool. - LP$	$L \leq 50$	2.19	1.18	1.12
	$L \leq 250$	1.57	0.54	0.62
	$L > 250$	1.57	0.51	0.58
	avg	1.85	0.83	0.84

Note: a: $y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + (1 - \mu^W) s^j_{Lt} l_{it} - \frac{1}{\eta} q^j_t + \tilde{a}_{it} + \tilde{\tilde{u}}_{it}$.
b: $y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + \gamma_L l_{it} - \frac{1}{\eta} q^j_t + \tilde{a}_{it} + \tilde{\tilde{u}}_{it}$.
c: $y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + \gamma_L l_{it} + a_{it} + u_{it}$

Table 2.4: Within estimates of μ_j, μ_{it}^W and ϕ_{it} for 21 industries

	a						*b*		
Sector j	$\widehat{\mu}_j$		$\widehat{\mu}^{Wj}$		$\widehat{\phi}^j$		$\widehat{\mu}_j$		N.obs
$j = 15$	1.141	(0.034)	0.515	(0.017)	0.381	(0.002)	1.173	(0.034)	7750
17	1.224	(0.063)	0.374	(0.043)	0.439	(0.004)	1.161	(0.053)	1581
18	1.235	(0.108)	0.592	(0.065)	0.611	(0.009)	1.297	(0.113)	396
19	1.065	(0.116)	0.911	(0.079)	0.692	(0.015)	1.084	(0.118)	347
20	1.665	(0.104)	0.647	(0.027)	0.645	(0.004)	1.292	(0.052)	1969
21	1.077	(0.042)	0.426	(0.027)	0.425	(0.003)	0.898	(0.027)	2500
22	1.253	(0.040)	0.504	(0.018)	0.442	(0.002)	1.085	(0.028)	6255
23	1.322	(0.376)	0.679	(0.176)	0.271	(0.002)	1.275	(0.352)	193
24	1.045	(0.036)	0.363	(0.030)	0.266	(0.002)	1.086	(0.038)	3885
25	2.730	(0.562)	0.436	(0.029)	0.456	(0.003)	1.439	(0.149)	3202
26	1.112	(0.068)	0.787	(0.027)	0.530	(0.005)	1.098	(0.064)	2704
27	0.940	(0.045)	0.883	(0.028)	0.578	(0.005)	0.857	(0.033)	1240
28	2.553	(0.248)	0.582	(0.016)	0.559	(0.002)	1.276	(0.054)	9885
29	1.241	(0.031)	0.602	(0.015)	0.600	(0.002)	1.036	(0.019)	8554
30	1.124	(0.092)	0.890	(0.077)	0.742	(0.002)	1.067	(0.079)	180
31	1.074	(0.029)	0.568	(0.034)	0.492	(0.005)	0.999	(0.024)	1808
32	1.122	(0.093)	0.437	(0.094)	0.609	(0.009)	0.998	(0.061)	398
33	1.055	(0.025)	0.511	(0.025)	0.522	(0.003)	0.993	(0.021)	2062
34	1.006	(0.031)	0.565	(0.040)	0.454	(0.005)	0.973	(0.027)	1574
35	1.091	(0.048)	0.788	(0.035)	0.664	(0.006)	1.019	(0.039)	1523
36	1.170	(0.021)	0.439	(0.030)	0.499	(0.003)	1.063	(0.016)	2666
avg	1.297		0.593		0.518		1.103		
$\rho_{\widehat{\mu}_j, \widehat{\mu}^{Wj}} = -0.188$	$\rho_{\widehat{\mu}^j, \widehat{\phi}^j} = -0.014$		$\rho_{\widehat{\phi}^j, \widehat{\mu}^{Wj}} = 0.694^{***}$						

Note: Estimating equation: $y_{it} = \gamma_0 + \gamma_K^j k_{it} + \gamma_M^j m_{it} + (1 - \mu^{Wj}) s_{Lt}^{\tilde{j}} l_{it} - \frac{1}{\eta_j} q_t^{\tilde{j}} + \tilde{a}_{it} + \tilde{\tilde{u}}_{it}$.
Standard errors in parentheses; sample period 1989-2008.

Table 2.5: IV estimates of μ, μ_{it}^W and ϕ_{it} for 21 industries

	a						*b*		
Sector j	$\widehat{\mu}_j$		$\widehat{\mu}^{Wj}$		$\widehat{\phi}^j$		$\widehat{\mu}_j$		N.obs
$j = 15$	1.511	(0.139)	0.208	(0.051)	0.129	(0.001)	1.496	(0.142)	7750
17	1.116	(0.090)	0.082	(0.104)	0.133	(0.004)	1.189	(0.120)	1581
18	1.146	(0.131)	0.457	(0.143)	0.698	(0.029)	1.373	(0.265)	396
19	0.899	(0.098)	0.034	(0.084)	0.073	(0.009)	0.921	(0.098)	347
20	3.380	(0.862)	0.114	(0.061)	0.195	(0.004)	1.272	(0.070)	1969
21	1.159	(0.131)	0.056	(0.094)	0.085	(0.002)	1.045	(0.087)	2500
22	1.197	(0.083)	0.064	(0.057)	0.066	(0.001)	1.053	(0.053)	6255
23	1.096	(0.090)	0.452	(0.150)	0.160	(0.008)	1.632	(0.465)	193
24	1.114	(0.095)	0.347	(0.079)	0.225	(0.003)	1.230	(0.116)	3885
25	1.169	(0.140)	0.246	(0.136)	0.343	(0.004)	1.396	(0.318)	3202
26	0.974	(0.116)	0.230	(0.078)	0.247	(0.003)	1.042	(0.117)	2704
27	0.968	(0.083)	0.546	(0.099)	0.439	(0.004)	0.956	(0.064)	1240
28	1.408	(0.115)	0.149	(0.047)	0.215	(0.002)	0.962	(0.048)	9885
29	1.483	(0.097)	0.094	(0.040)	0.076	(0.001)	1.131	(0.041)	8554
30	1.108	(0.144)	0.095	(0.223)	0.311	(0.019)	0.940	(0.108)	180
31	1.271	(0.104)	0.070	(0.132)	0.103	(0.005)	1.078	(0.061)	1808
32	1.093	(0.159)	0.356	(0.131)	0.859	(0.064)	0.963	(0.127)	398
33	1.030	(0.051)	0.088	(0.067)	0.240	(0.008)	0.958	(0.038)	2062
34	0.953	(0.033)	0.142	(0.101)	0.097	(0.001)	1.026	(0.037)	1574
35	1.066	(0.092)	0.102	(0.099)	0.253	(0.010)	0.987	(0.059)	1523
36	1.179	(0.037)	0.139	(0.100)	0.160	(0.002)	1.027	(0.024)	2666
avg	1.253		0.194		0.243		1.128		

$\rho_{\widehat{\mu}^j,\widehat{\mu}^{Wj}} = -0.173^{***}$ $\rho_{\widehat{\mu}^j,\widehat{\phi}^j} = -0.054^{***}$ $\rho_{\widehat{\phi}^j,\widehat{\mu}^{Wj}} = 0.726^{***}$

Note: Estimating equation: $y_{it} = \gamma_0 + \gamma_K^j k_{it} + \gamma_M^j m_{it} + (1 - \mu^{Wj}) s_{Lt}^{\tilde{j}} l_{it} - \frac{1}{\eta^j} q_t^{\tilde{j}} + \tilde{a}_{it} + \tilde{\tilde{u}}_{it}$.
Standard errors in parentheses; sample period 1989-2008.

Table 2.6: Wool.–LP estimates of μ, μ^W_{it} and ϕ_{it} for 21 industries

	a						*b*		
Sector j	$\widehat{\mu}_j$		$\widehat{\mu}^{Wj}$		$\widehat{\phi}^j$		$\widehat{\mu}_j$		N.obs
$j=15$	1.203	(0.069)	0.228	(0.043)	0.139	(0.001)	1.173	(0.064)	7750
17	1.030	(0.070)	0.124	(0.087)	0.189	(0.006)	1.027	(0.068)	1581
18	1.236	(0.166)	0.457	(0.143)	0.698	(0.029)	1.288	(0.172)	396
19	0.869	(0.085)	0.006	(0.103)	0.013	(0.002)	0.917	(0.082)	347
20	3.446	(0.843)	0.105	(0.060)	0.183	(0.004)	1.218	(0.063)	1969
21	1.015	(0.082)	0.074	(0.068)	0.109	(0.003)	0.986	(0.078)	2500
22	1.209	(0.081)	0.154	(0.052)	0.146	(0.002)	1.068	(0.060)	6255
23	0.952	(0.054)	0.288	(0.199)	0.109	(0.006)	–	–	193
24	1.056	(0.092)	0.373	(0.075)	0.238	(0.003)	1.095	(0.098)	3885
25	1.255	(0.177)	0.235	(0.120)	0.333	(0.004)	1.066	(0.123)	3202
26	0.954	(0.106)	0.253	(0.086)	0.265	(0.003)	0.974	(0.110)	2704
27	0.901	(0.065)	0.406	(0.123)	0.368	(0.004)	0.885	(0.057)	1240
28	1.445	(0.111)	0.146	(0.041)	0.211	(0.002)	0.979	(0.044)	9885
29	1.362	(0.075)	0.087	(0.041)	0.071	(0.001)	1.098	(0.040)	8554
30	1.263	(0.113)	0.025	(0.194)	0.105	(0.008)	0.996	(0.050)	180
31	1.267	(0.095)	0.149	(0.096)	0.197	(0.008)	1.065	(0.061)	1808
32	0.972	(0.143)	0.363	(0.125)	0.862	(0.063)	0.853	(0.105)	398
33	1.077	(0.056)	0.100	(0.077)	0.264	(0.009)	1.000	(0.046)	2062
34	0.938	(0.030)	0.000	(0.115)	0.000	(0.000)	0.972	(0.029)	1574
35	1.046	(0.079)	0.120	(0.092)	0.286	(0.010)	0.934	(0.055)	1523
36	1.105	(0.032)	0.042	(0.091)	0.054	(0.001)	0.994	(0.024)	2666
avg	1.219		0.178		0.230		1.029		

$\rho_{\widehat{\mu}^j,\widehat{\mu}^{Wj}} = -0.199^{***}$ $\rho_{\widehat{\mu}_j,\widehat{\phi}^j} = -0.289^{***}$ $\rho_{\widehat{\phi}^j,\widehat{\mu}^{Wj}} = 0.782^{***}$

Note: Estimating equation: $y_{it} = \gamma_0 + \gamma^j_K k_{it} + \gamma^j_M m_{it} + (1-\mu^{Wj}) s^{\tilde{j}}_{Lt} l_{it} - \frac{1}{\eta^j} q^{\tilde{j}}_t + \tilde{a}_{it} + \tilde{\tilde{u}}_{it}$.
Standard errors in parentheses; sample period 1989-2008.

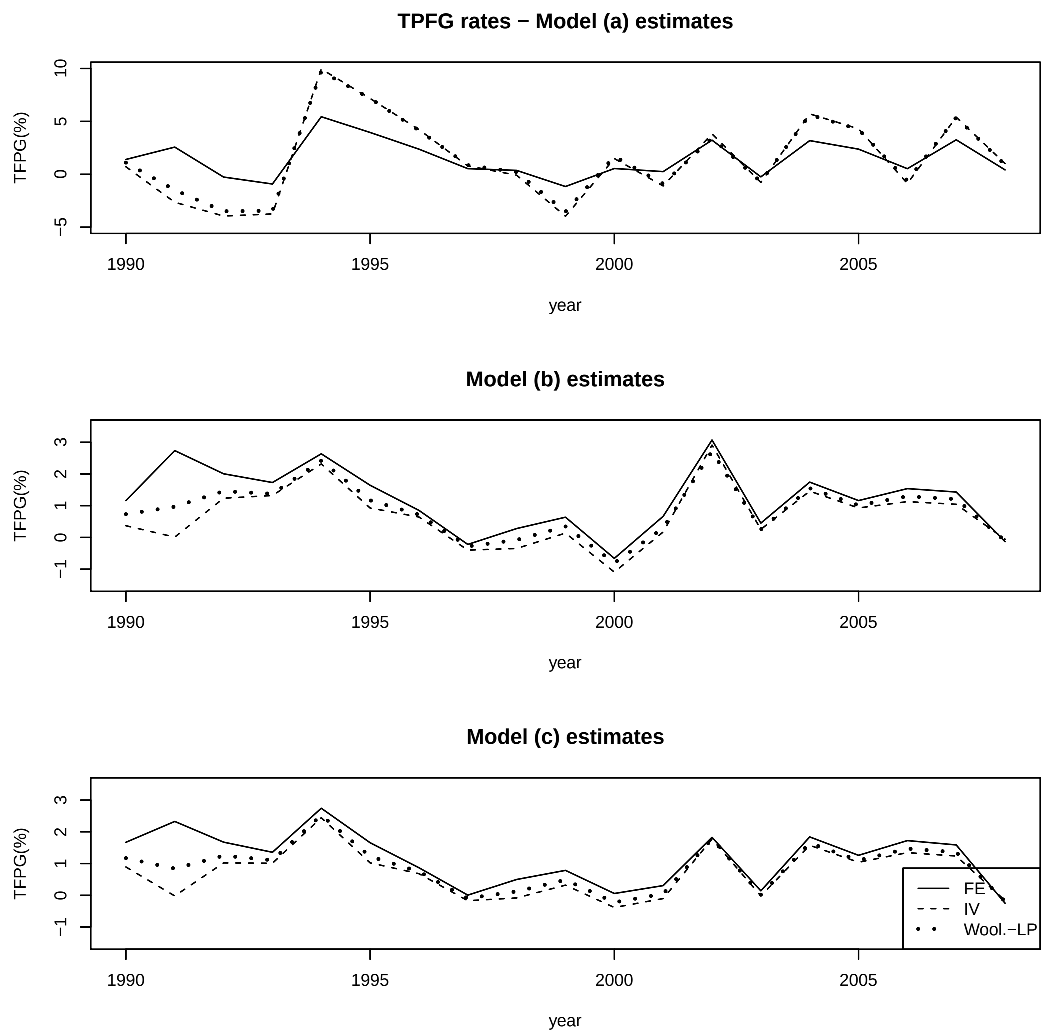
Table 2.7: Weighted average TFP growth rates per sector

	$TFPG(a)$		$TFPG(b)$	
sector	IV	Wool.-LP	IV	Wool.-LP
15	0.49	1.00	-0.44	0.13
17	1.09	0.81	1.11	0.80
18	2.62	3.35	2.88	2.55
19	3.16	2.23	1.27	1.02
20	1.64	1.47	-0.69	-0.85
21	0.69	0.87	0.39	0.53
22	0.34	0.37	0.16	0.21
23	–	–	–	–
24	0.98	1.26	0.47	0.90
25	1.99	2.15	1.76	1.29
26	0.79	0.89	0.49	0.59
27	1.39	1.92	0.80	1.03
28	2.84	2.74	1.06	1.07
29	3.24	2.91	1.00	0.91
30	3.82	5.75	1.08	1.44
31	3.13	3.34	1.72	1.62
32	2.05	3.64	2.48	3.02
33	2.70	2.70	1.05	1.05
34	2.45	2.28	1.14	1.33
35	3.83	3.40	1.43	1.29
36	1.12	0.80	0.32	0.15
avg	1.90	1.91	1.28	0.73

Table 2.8: Correlation between TFP growth and structural parameters

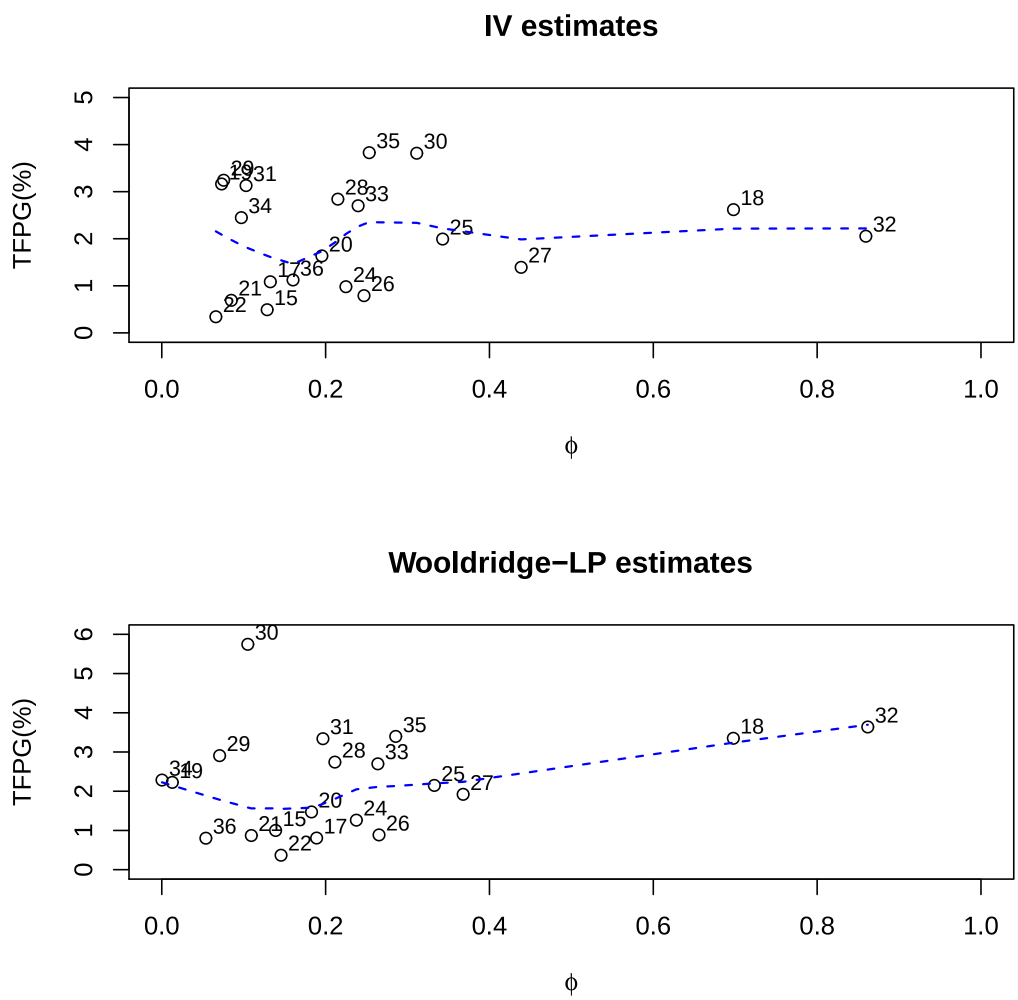
ρ	TFPG(IV)	TFPG(Wool.-LP)
μ_j	-0.0833	0.0857
μ^{jW}	-0.0195	0.0779
ϕ^j	0.1519	0.1636

Figure 2.1: TFP Growth by estimation approach



Note: Sample period 1989-2008; dependent variable: log. gross deflated output y_{it} $TFPG_{it} = \pi^j_{it}\Delta\hat{a}_{it} \equiv \pi^j_{it}\Delta\left[\hat{\mu}_j\left(y_{it} - \left(\hat{\gamma}_K k_{it} + \hat{\gamma}_M m_{it} + (1-\hat{\mu}^W)s_{iLt}l_{it} - \frac{1}{\hat{\eta}_j}q^j_t\right)\right)\right]$, where $\pi^j_{it} = L_{it}/\sum_i L^j_{it}$

Figure 2.2: TFP Growth per sector



Note: Weighted average TFPG over the period 1990–2008 from model a (labor and market imperfections).

Table 2.9: Averages per NACE 2-digit sector code

Sector	y_{it}	l_{it}	m_{it}	k_{it}	s_{iLt}	s_{iMt}	s_{iKt}	Q_t^J	$N.obs.(N.firms)$
15	9.311	4.185	8.904	6.089	0.200	0.688	0.055	0.991	7750(884)
17	8.817	4.037	8.254	5.562	0.273	0.618	0.043	1.047	1581(160)
18	7.931	3.464	7.325	4.150	0.294	0.596	0.029	1.297	396(56)
19	7.387	2.914	6.803	3.885	0.256	0.596	0.034	0.955	347(62)
20	8.609	3.890	8.204	5.157	0.264	0.651	0.033	0.911	1969(219)
21	9.468	4.437	9.002	6.562	0.241	0.633	0.058	0.929	2500(203)
22	8.522	3.821	7.921	5.457	0.312	0.560	0.056	1.032	6255(774)
23	10.264	4.566	9.820	7.598	0.129	0.735	0.060	1.059	193(23)
24	9.741	4.362	9.214	6.630	0.192	0.669	0.052	1.004	3885(419)
25	9.166	4.217	8.654	6.123	0.245	0.631	0.055	0.907	3202(321)
26	8.974	3.989	8.401	5.955	0.247	0.599	0.057	0.947	2704(275)
27	9.284	4.310	8.764	6.069	0.244	0.643	0.046	0.947	1240(125)
28	8.548	3.886	7.974	5.113	0.300	0.588	0.041	0.886	9885(1116)
29	8.853	4.110	8.306	5.283	0.303	0.595	0.033	0.881	8554(907)
30	9.396	4.428	8.895	5.375	0.286	0.634	0.034	1.128	180(23)
31	8.729	3.939	8.141	5.030	0.289	0.597	0.031	1.045	1808(211)
32	8.451	3.758	7.811	4.991	0.301	0.583	0.050	1.013	398(74)
33	8.454	3.905	7.692	4.874	0.356	0.514	0.037	1.040	2062(269)
34	9.009	4.174	8.576	5.240	0.248	0.670	0.029	0.932	1574(163)
35	8.889	4.013	8.395	4.900	0.270	0.643	0.026	0.922	1523(187)
36	8.476	3.918	7.931	5.003	0.301	0.603	0.035	0.944	2666(305)
Total	8.892	4.046	8.358	5.555	0.270	0.615	0.044	0.953	60672(6727)

Note: Averages per sector through the sample period 1989-2005. NACE two-digit codes: food products, beverages and tobacco (15-16); textile and leather products (17-19); wood (20); paper, paper products, publishing and printing (21-22); coke, refined petroleum products and nuclear fuel (23); chemicals and chemical products (24); rubber and plastic products (25); other non-metallic mineral products (26); basis metals and fabricated metal products (27-28); machinery and equipment n.e.c. (29); electrical and optical equipment (30-33); transport equipment (34-35); other manufacturing activities (36).

Chapter 3

Multilevel Heterogeneity of R&D Cooperation Determinants

Abstract. Using data from the 2006 edition of the Community Innovation Survey (CIS) for the Netherlands, we propose a methodology to study the effect of firm-level characteristics on the propensity to undertake a research collaborative agreement. In particular, we show that controlling for a richer variance structure, that describes the expected correlation present among firms within a particular sector, leads to the correct innovation policy impact evaluation. The suggested multilevel design, with firms nested in sectors, proves to be the gauge of the effectiveness of public innovation policies. Moreover, such a hierarchical framework can be generalized allowing for clustering at higher levels, such as sectors or geographical areas. In line with the literature on R&D cooperation determinants, our results confirm the importance of technological spillovers, risk and cost sharing rationales, firm's size, and type of innovative activity in influencing the decision to engage in different sorts of research alliances. Moreover, we find that, in addition to the firm-level heterogeneity in cooperation strategies, there is correlation between firms within each sector. In fact, enterprises within the same industry could face similar market conditions, such as the level of competition. Therefore, R&D cooperation decisions, as well as innovative production, are firm-level processes, where a strong sectoral specificity exists.

3.1 Introduction

Given the increasing market pressures, a firm that wants to survive must not only be innovative, but also ready to face shorter business cycles,[17] and prompt to meet a more dynamic demand. Phenomena such as knowledge outsourcing and networking are at the core of entrepreneurial actions. In particular, firms decide to collaborate in research for various economic reasons. In fact, a research alliance could aim at strengthening the member firms' core competencies, so as to reach for new markets, or it may be a strategic decision to access complementary knowledge in order to compensate for the absence of internal competencies or to reduce the costs associated with knowledge spillovers. The theoretical literature on cooperative R&D points at the internalization of the technological spillovers as the main rationale behind the decision to cooperate (Katz, 1986; d'Aspremont and Jacquemin, 1988; Kamien et al., 1992). Parallel to the theory, the empirical literature confirms the relevance of such spillovers in influencing the choice of cooperating, and extends the analysis on R&D cooperation determinants by distinguishing between incoming and outgoing spillovers (Cassiman and Veugelers, 2002, 2006; Lopez, 2008). Moreover, several authors focus on the analysis of the heterogeneity in the determinants of innovating firms' decisions to engage in R&D cooperation (Kaiser, 2002; Belderbos et al., 2004b,a). These studies explore the differences in the factors affecting the firms' probability to establish different types of cooperation, namely horizontal cooperation (with competitors), vertical cooperation (with suppliers or customers), and institutional cooperation (with universities and research institutes). Furthermore, Belderbos et al. (2004a) relax the assumption on the independence among different cooperation strategies, accounting for possible correlations between the strategies that could be due to technological complementarities. However, the existing literature on the determinants of R&D cooperation with different R&D partners overlooks the role of sectoral specificities (for example, sector-specific physical assets) in influencing the expected correlation among the different cooperative strategies present

[17]Recent research employing spectral analysis has confirmed the presence of sinusoidal-like cycles (called Kondratiev) in the world GDP dynamics at an acceptable level of statistical significance. Korotayev and Tsirel (2010) detected shorter (on average 17 years) business cycles, approximately one third of the Kondratiev cycles.

among firms within a particular sector.[18] As a matter of fact, horizontal R&D cooperations are likely to be formed within the same sector as it will lead to *collective efficiencies* (Schmitz, 1999) in the form of reduced transaction costs and accelerated innovation rates through a greater market access. These collective efficiencies are of particular interest from a policy perspective. In fact, they can be interpreted as the multiplier effect of an innovation policy: the increased innovative intensity of one company or several companies multiplies the economic benefits in a given sector by helping to drive the innovative intensity of other business entities. This type of externality is demand-driven, in the sense that the private and public for innovation can be stimulated by new or growing business enterprises, which enables their suppliers to grow as well.

Such collective interactive processes may derive from organizational proximities, and when organizational proximity arises between organizations connected by a relationship of either economic or financial dependence/interdependence, Kirat and Lung (1999) state that intra-sectoral links are liable to dominate inter-sectoral links.

Therefore, in this paper, we control for a richer variance structure that describes the expected correlation present among firms' cooperative choices within a particular sector. In particular, we analyze the correlation among different R&D alliances due to both firm- and sector-level heterogeneity, adopting a multivariate hierarchical logit model. The advantages of adopting a hierarchical structure (often referred to as multilevel, random or mixed, (see Goldstein, 1995; Hedeker and Gibbons, 1996)) are several. First, it allows us to assume and specify a more complex covariance structure. This means that we can fit a regression model to firm-level data, while accounting for unexplained variation among the sectors, aiming at capturing relevant features to explain the propensities to undertake a specific cooperative agreement. Second, there is no need to have a balanced design or equally spaced measurements, as the number of firms per sector is allowed to vary. Third, unlike the multivariate probit, the logit specification is not restricted to the normal distribution assumption for the individual effects, and

[18]Depending on the model assumptions, and compatibly with the data at hand, one could allow for a richer specification of the clusters, such as the geographical district, or the relevant markets. We limit ourselves to a frugal, yet general representation of a multilevel design in the context of research cooperation determinants.

its statistical fit is more accessible because of the odds-ratio interpretation of the logit coefficients. Its derivation is straightforward, and simulation of its choice probabilities is computationally simple.

Our approach departs from the one used to test for complementarities (Mohnen and Roller, 2005; Belderbos et al., 2006), as our main focus is not to estimate the degree of strategic complementarities or substitutabilities among firms' cooperation choices, but rather to model and estimate both individual and aggregate forms of externalities, in which the collective actions of a reference group affect an individual's choices. As pointed out by Mohnen and Roller (2005), innovation policies may have different impacts on the distinctive phases of the innovation process. As a matter of fact, while there could be firm-level policy externalities in the decision to collaborate in research, the innovative produce might well be affected by the aforementioned demand driven innovation policy externality.

Therefore, to explicitly take into account both firm- and sector-level externalities, and the different impact of innovation policy measures on the two phases of the innovative process, we divide our study in two stage. In the first stage we study the main drivers of undertaking a collaborative agreement with a research partner. In the second stage, we investigate the effects of innovation policies and R&D cooperation on innovative intensity.

Using data from the 2006 edition of the Community Innovation Survey (CIS2006) on 1,947 innovating Dutch firms operating in 15 manufacturing sectors, we analyze the firm- and sector-level heterogeneity of the determinants of R&D cooperations and of the production of innovative output determinants. To our knowledge, this paper is the first attempt to model the firms' determinants of R&D partner's choices and to assess the impact of public financial support to innovative output through a hierarchical model.

Our results confirm the hypothesis of a more structured correlation assumption. As a matter of fact, additional to the well documented firm-level heterogeneity (Belderbos et al., 2004a), we find evidence of sector-level heterogeneity in the variables explaining the probabilities to cooperate and in the factors affecting the level of innovative production. Moreover, controlling for public financial support for innovation activities at different levels of government, we show that when using the suggested multilevel approach the impact of public funding has a positive and significant sign, while, when omitting this nested framework, the policies have a poor effect on innovative turnover.

The remainder of this paper is organized as follows. In Section 2, we look at the existing theoretical and empirical literature on the R&D cooperation determinants to guide our own empirical analysis. Section 3 describes the model. In Section 4, we describe the data. Section 5 and 6 discuss the empirical results of the mixed models to describe both the propensity to undertake the different cooperation agreements, and the impact of the public funding on the firm-level innovative output. Section 7 summarizes.

3.2 Determinants of R&D cooperation: a review of firm-level and industry-level factors

The Industrial Organization (IO) literature has pointed to technological spillovers as one of the important factors influencing the firms' incentives to engage in cooperative R&D (Katz, 1986; d'Aspremont and Jacquemin, 1988; Kamien et al., 1992). R&D spillovers arise when new knowledge created by one firm is also beneficial to other firms. Theoretical studies suggest that a high level of R&D spillovers can increase the firms' probability of internalizing the spillovers by participating in R&D cooperation.

The relationship between spillovers and R&D cooperation has also been analyzed using empirical studies. Cassiman and Veugelers (2002) empirically explore the effects of knowledge flows on R&D cooperation on Belgian firms' decisions to enter into a cooperative R&D agreement, highlighting the distinction between the effect of knowledge flows into firms (*incoming spillovers*) and that of outbound knowledge flows (*appropriability*). Cassiman and Veugelers (2002) observe a significant relation between external information flows and the decision to cooperate in R&D. Indeed, the most important finding of their two-step probit model is that the probability of firms cooperating in R&D is higher when incoming spillovers are high and outgoing spillovers are low. Furthermore, cost-sharing is found to be an important motive for cooperation in R&D, while risk-sharing is not. The empirical model of Cassiman and Veugelers (2002) has been to some extent modified and applied to Spanish firms by Lopez (2008), who focuses on the role of cost and risk sharing of innovation projects for the cooperation decision, and finds evidence supporting the importance of cost/risk sharing as a hampering factor for the innovation process. Further, the paper of Lopez (2008) pays much attention to the endogeneity of the independent variables which in other papers are assumed to be endogenous a priori. The results confirm the conclusions of Cassiman and Veugelers (2002) that spillovers and appropriability play an important part in influencing R&D cooperation decisions, only when an adequate structure of endogeneity is chosen. Indeed, the hypothesis of exogeneity of incoming spillovers and appropriability is rejected, while the exogeneity of R&D intensity cannot be rejected.

In a study based on data of European firms, Hernán et al. (2003) follow a two-step procedure. In the first stage, the entire population that could potentially participate in a cooperative organization is considered. In this first stage, it is possible to measure the effect of the relevant firm characteristics that influence a Research Joint Venture (RJV) formation. In the second stage, the focus is on firms that are known to have a higher probability to participate in cooperative R&D projects. Using a large database of firms from almost twenty European countries, Hernán et al. (2003) find that, contrary to what has been found by Cassiman and Veugelers (2002) and more in line with the aforementioned IO theory, patents' effectiveness and, therefore, the level of appropriability, reduces R&D cooperations. Moreover, among individual firm characteristics, firm size, and previous participation experience increase the likelihood of participating. Industry-level characteristics are also significant, especially R&D intensity. RJVs are also more likely in more concentrated industries where technological knowledge diffuses rapidly. Therefore, a minimum level of industry concentration is needed for RJVs to be formed. One possible explanation for the significance of firm size is that EU programs favor large partnerships, which may be more costly to manage. An important characteristic of their analysis is that they use a large control group that is representative for the whole population of European firms. With respect to differences between countries, Hernán et al. (2003) find that mainly firms in smaller countries participate in projects funded by the EU; according to the authors this is because firms in large countries can find partners in their own country more easily. Spillovers are measured at the industry level and are proxied by the average number of months before the diffusion of an innovation in the industry and the effectiveness of patents in the industry, both based on previous analyses. Problems of endogeneity are dealt with by lagging all time-dependent right-hand-side variables by two years.

New and sharper results dealing with the relationship between R&D cooperation, spillovers, and productivity appear in Belderbos et al. (2004a), who construct a multivariate cross-sectional probit model to explore differences in the determinants of innovating firms' decisions to participate in four distinct types of partner specific innovation strategies (cooperation with competitors, suppliers, customers, and universities and research institutes). With a large Dutch dataset (627 firms with R&D cooperation of some type), evidence of a positive impact of R&D cooperation on labor productivity growth is found, but with distinct differences depending on (the combination of) cooperation types. Competitor and supplier cooperations seem to have the most positively significant impact on productivity growth. The results for the other variables show that incoming spillovers and R&D intensity are statistically significant in explaining R&D cooperation with firms from the same industry.

Busom and Fernández-Ribas (2008) find that the formation of inter-firm alliances is likely to vary among firms and can be due to a large number of reasons, such as the nature and the scope of the R&D projects. In particular, the authors show that if a firm's aim in a cooperative agreement is to share complementary technology, it will tend to cooperate with heterogeneous partners (heterogeneous in knowledge assets, market scope, location, and product range), while, when the motivation for cooperation is based on internalizing outgoing spillovers or on increasing market power, symmetric partnerships (i.e., horizontal or vertical) are more likely.

A less developed strand of literature deals with the *innovativeness* of firms and their propensity to cooperate. As a matter of fact, some R&D alliances may differ in the degree of technological effort required. Indeed, we can identify two types of innovations that R&D partnerships may develop: radical and incremental.[19] A radical innovation is a product, process, or service offering "significant

[19]The labels *radical* and *incremental* belong mostly to the managerial literature (see Dewar and Dutton (1986); Henderson (1993)).

improvements in performance or cost that transform existing markets or create new ones" (Leifer et al., 2001). On the other hand, incremental innovations are based on minor changes or improvements in the current technology. Nonetheless, similar concepts are implicitly used in the economic literature. IO theorists, as for example Reinganum (1983), use terms such as *drastic innovations* to describe those changes in technology that determine a decrease in costs such that the new equilibrium price lies below the pre-innovation cost and consequently turn the innovator into a monopolist. On the other hand, non-drastic or *gradual* innovations only introduce costs asymmetries that do not transform the market into a monopoly. Tether (2002) observed that true or radical innovators cooperate more than those who introduce only imitative innovations.

None of these studies, however, controls for the likely multilevel structure of the data. As a matter of fact, data may occur in clusters, such as sectors in which firms are nested. One approach to modeling such type of data includes random effects for subjects (firms) or clusters (sectors) into account. This provides a mechanism of accounting for certain correlation structures among the clustered observations.

3.3 Model specification

To investigate the relationship between the factors driving the propensity to collaborate with different research partners, we assume a hierarchical structure of the model specification. In particular, we adopt a *multivariate mixed logit model* (multi-response Generalized Linear Mixed Model– GLMM, (Hedeker and Gibbons, 1996)).[20] Using the terminology of multilevel analysis, let i denote the level-1 units (nested observations, i.e., firms) and let j denote the level-2 units (subjects, i.e., sectors). Assume there are $j = 1, \dots, J$ sectors and $i = 1, \dots, N$ firms. The total number of firms is given by $N = \sum_{j=1}^{J} n_j$, where n_j is the number of firms nested within each sector (level-2 unit). Each firm i is faced with $c = 1, \dots, C$ different choices of cooperation strategies.

[20] The class of mixed logit models is a highly flexible as it can approximate any random utility model (Train, 2009). The results we present can be generalized and extended to panel data.

Let us define the utility that firm i in sector j obtains from choosing cooperation c as

$$\begin{aligned} U_{ij}^c &= V_{ij}^c + \epsilon_{ij}^c \qquad \text{where} \\ V_{ij}^c &= \beta_0^c + x_{ij}'\beta^c + z_{ij}'\alpha_j^c + \gamma_i^c. \end{aligned} \tag{3.1}$$

β_0^c and β^c are the intercept and the vector of category-specific fixed effects, respectively, x_{ij} and z_{ij} are vectors of observed variables. The error term ϵ_{ij}^c is assumed to be independently, identically extreme value distributed. The firm- and sector-level random coefficients, γ_i^c and $\alpha_j^c \equiv (\alpha_{1j}, \ldots, \alpha_{qj}^c)'$ are assumed to be normally distributed,

$$\alpha_j^c \sim N_q(0, W) \qquad \text{and} \qquad \gamma_i^c \sim N(0, r),$$

where q is the number of random effects included in the model. The variances of the firm- and sector-level random components, r and $W = diag(w_1, \ldots, w_q)$, respectively, are assumed to be invariant to cooperation choice c. We define the random effects for all sectors as $\alpha^c \equiv ((\alpha_1^c)', \ldots, (\alpha_J^c)')'$, for all firms as $\gamma^c \equiv (\gamma_1^c, \ldots, \gamma_N^c)'$, and for all c cooperation strategies as $\alpha \equiv \big((\alpha^1)', \ldots, (\alpha^C)'\big)'$, and $\gamma \equiv \big((\gamma^1)', \ldots, (\gamma^C)'\big)'$. We then assume that the vector of firm- and sector level random intercepts and slopes, γ and α have the following covariance structure:

$$\gamma \sim N(\mathbf{0}, \mathbf{G_1}) \qquad \text{and} \qquad \boldsymbol{\alpha} \sim N(\mathbf{0}, \mathbf{G_2}).$$

$\mathbf{G_1}$ and $\mathbf{G_2}$ are defined as the Kronecker product between matrices $\mathbf{A_1}$, and $\mathbf{A_2}$, and $\mathbf{V_1}$, and $\mathbf{V_2}$, i.e., $\mathbf{G_1} = \mathbf{V_1} \otimes \mathbf{A_1}$, and $\mathbf{G_2} = \mathbf{V_2} \otimes \mathbf{A_2}$ where

$$\mathbf{V_1} = \begin{pmatrix} \varsigma_{11}^2 & \varsigma_{12} & \cdots & \varsigma_{1c} \\ \varsigma_{21} & \varsigma_{22}^2 & \cdots & \varsigma_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ \varsigma_{c1} & \varsigma_{c2} & \cdots & \varsigma_{cc}^2 \end{pmatrix} \qquad \text{and} \qquad \mathbf{V_2} = \begin{pmatrix} \sigma_{11}^2 & \sigma_{12} & \cdots & \sigma_{1c} \\ \sigma_{21} & \sigma_{22}^2 & \cdots & \sigma_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{c1} & \sigma_{c2} & \cdots & \sigma_{cc}^2 \end{pmatrix} \tag{3.2}$$

are the firm- and sector-level cooperation strategy-specific covariance matrices with elements $\varsigma_{c\tilde{c}} = \varsigma_{\tilde{c}c}$, and $\sigma_{c\tilde{c}} = \sigma_{\tilde{c}c}$, for $c \neq \tilde{c}$. In our application, these are

4×4 matrices, as we consider 4 types of R&D cooperation ($C = 4$), and where

$$\mathbf{A_1} = diag(r_1, \ldots, r_N) \qquad \text{and} \qquad \mathbf{A_2} = diag(W_1, \ldots, W_J).$$

The matrices $W_1, \ldots, W_J$ have dimension $q \times q$, so that the dimension of $\mathbf{A_2}$ is $qJ \times qJ$.[21] $\mathbf{G_1}$ and $\mathbf{G_2}$ are block matrices of dimension $4N \times 4N$ and $4qJ \times 4qJ$, respectively. The model specification can be generalized so as to take into account for both random coefficients and heteroskedasticity, by using the following specification for the random intercept and coefficients:

$$\alpha_j^c \sim N_q(0, W_j) \qquad \text{and} \qquad \gamma_i^c \sim N(0, r_i).$$

The variances of the firm- and sector-level random components, r_i and $W_j = diag(w_{1j}, \ldots, w_{qj})$, would then measure the degree of heterogeneity of each firm, nested in each sector. In this paper, the only source of heteroskedasticity which is explicitly taken into account is the one deriving from different cooperation alternatives, $\varsigma_{c\tilde{c}}$ and $\sigma_{c\tilde{c}}$.

The regression parameters are collected in the vector $\theta = ((\beta_0, (\beta^c)', \alpha', \gamma')$. Given the extreme value distribution assumption of the error term ϵ_{ij}^c, the likelihood for firm i in sector j of the binary decision to cooperate partner c (independently from other alternatives) follows a logistic distribution, and can be written in a closed form expression:

$$f_{ij}^c(y_{ij}^c|\theta) = \frac{\exp(V_{ij}^c)}{1 + \exp(V_{ij}^c)}, \tag{3.3}$$

where $y_{ij}^c \in \{0, 1\}$ is the observed research cooperation choice. Assuming *conditional independence* of firm's choice probabilities given the covariates and the random effects, we can write the unconditional marginal probability[22] of the

[21] If we set the error component $z_{ij}'\alpha_j^c = d_{ij}'\alpha_j^c$, where d_{ij} is a dummy variable that takes the value 1 if firm i is nested in sector j and zero otherwise, α_j^c is reduced to a category-specific random intercept. In such a case $W_j = w_j$ and $\mathbf{A_2} = diag(w_1, \ldots, w_J)$ would be a simple diagonal matrix of dimension $J \times J$.

[22] Coull and Agresti (2000) derive a multivariate Binomial logit-normal distribution, where the c responses $Y_i = (Y_{i1}, \ldots, Y_{ic})$ with index vector $m_i = (m_{i1}, \ldots, m_{ic})$ are assumed to be independent binomial distributions, with success parameter vector π_i. Then the multivariate Binomial logit-normal model is expressed by incorporating a random effect, such that

response block matrix, $\mathbf{Y} \equiv \left[y_{ij}^{c}\right]_{4\times N\times J}$ as

$$L(\mathbf{Y}|\mathbf{G_1},\mathbf{G_2}) = \int\int \prod_j \prod_c \prod_i f_{ij}^{c}(y_{ij}^{c}|\theta)\pi_1(\gamma|\mathbf{G_1})\pi_2(\boldsymbol{\alpha}|\mathbf{G_2})d\gamma d\boldsymbol{\alpha}. \tag{3.5}$$

The prior densities $\pi_1(\gamma|\mathbf{G_1})$ and $\pi_2(\boldsymbol{\alpha}|\mathbf{G_2})$ (also called mixing distributions) are assumed to be independent (see the Appendix for a description of a Bayesian Multivariate Mixed Logit Model).

In the next subsections, we describe the data, and how we construct the firm- and sector-level characteristics influencing the choice among the different research partners.

3.4 Data

The data used for the present study corresponds to the 2006 edition of the Community Innovation Survey[23] (also referred to as CIS2006), carried out by Statistics Netherlands.

The Dutch CIS2006 collected data on product and process innovation, as well as organisational and marketing innovation during the three-year period 2004 to 2006. The total number of manufacturing firms participating to the survey was 1,929.

$logit(\pi_i) = X_i\beta + z_i$. where X_i is a $c \times p$ covariate matrix and z_i is a $c \times 1$ vector of random effects and is distributed as a multivariate normal distribution with mean vector $\mathbf{0}$ and covariance matrix Σ. Then the probability density function of y is written as

$$p(y;\pi,m,\Sigma) = \int_{[0,1]^c} f_B(y|\pi,m) f_N(z;\Sigma)dz \tag{3.4}$$

where $f_B(y|\pi,m)$ denotes the binomial probability mass function with m trials and success probability π and $f_N(z;\Sigma)$ denotes the multivariate normal density function of z.

[23]The Community Innovation Surveys are designed to provide an extensive description of the general structure of innovative activities at the country and industry levels. Within the guidelines of the OSLO Manual on performing innovation surveys (OECD, 1997), information about innovation activities is collected.

3.4.1 Dependent Variables

The dependent variables of the model are dummy variables equal to one if the firm, during the three years 2004 to 2006, actively participated with other enterprises or non-commercial institutions on innovation activities. In particular, as in Belderbos et al. (2004a), we consider four R&D partnerships, namely with suppliers, clients, competitors, or research institutes and/or universities (institutional cooperation).

Cooperation networks can be further distinguished in two types of cooperative behavior. The first is based on the synergies obtained by combining complementary assets. This combination of resources enables a more complete or intense use of the different types of assets possessed by each firm. One of the most important of these complementary cooperation agreements is the *vertical* or *supply-chain* cooperation, in which the company cooperates with its customers and/or suppliers (Tether, 2002).

The second rationale which characterizes the other type of cooperative agreement consists in competitive positioning (Cassiman and Veugelers, 2002), i.e., seeking market power. This type of cooperation, also known as *horizontal* R&D cooperation tends to form matches between competing firms that might have similar needs in terms of product or process development, looking for resources of the same type (technological, human, and so on). A summary of the percentages of firms adopting the four types of R&D cooperation is presented in Table 3.1. Vertical cooperation, Co_{ij}^{vert}, is the most frequent type of agreement, as 27% of firms decide to engage in either a collaboration with suppliers, Co_{ij}^{supp} (25%), or with customers, Co_{ij}^{cust} (19%). The last column of the table reports the number of firms undertaking the various collaborations. R&D cooperation with other competing enterprises is the least frequently observed choice (7% of the total sample). Collaboration with universities and public institution of research is chosen by 278 out of 1927 firms.

Table 3.2 reports percentages (and numbers) of firms undertaking at least one cooperation alliance that maintain cooperative R&D agreements with the same or different partners. It is interesting to note that, even though most of the R&D cooperations are formed between the same type of partner (the percentages on

Table 3.1: Summary statistics

Cooperation type	% of coop. firms	sd	N.of firms
Co_{ij}	0.285	0.451	549
Co_{ij}^{supp}	0.250	0.433	481
Co_{ij}^{cust}	0.186	0.389	359
Co_{ij}^{comp}	0.075	0.263	144
Co_{ij}^{inst}	0.144	0.351	278
Co_{ij}^{vert}	0.273	0.446	526

Note: Average percentages and number of firms engaging in different types of cooperation agreements: suppliers, clients, competitors, or public research institutions. Vertical cooperation is here defined as the cooperation with suppliers or customers ($Co_{ij}^{vert} \equiv Co_{ij}^{supp} \cup Co_{ij}^{comp}$).

Table 3.2: Cooperative R&D agreement combinations

	Suppliers	Customers	Competitors	Institutional
Suppliers	81.16(547)	53.26(359)	23.89(161)	50.15(338)
Clients or customers	-	60.98(411)	20.77(140)	40.65(274)
Competitors	-	-	26.71(180)	21.07(142)
Institutional	-	-	-	59.94(404)

Note: R&D partner choices composition of firms with at least one cooperative R&D agreement. Absolute frequencies in parentheses.

the diagonal are larger than the off-diagonal ones), a large share of firms is exchanging knowledge with other sorts of collaborators. For example, 338 firms which cooperate with suppliers have agreements also with research institutions, while only 161 firms cooperating with suppliers maintain agreements with competitors.

3.4.2 Independent Variables

Following the existing theoretical and empirical work, we propose four sets of explanatory variables related to firms' characteristics, obstacles to innovation that the firm should overcome, the existence of public funding to encourage R&D, and sectoral characteristics within which the firm operates. With regard to firms' characteristics, firm-level knowledge inflows, also defined as incoming spillovers (Cassiman and Veugelers, 2002), are derived from the scores of importance of

Table 3.3: Summary statistics

variable	mean	sd	min	max
In_{ij}	0.572	–	0	1
rad_{ij}	0.317	–	0	1
inc_{ij}	0.336	–	0	1
$spill_{ij}$	0.161	0.255	0.000	1.000
$size_{ij}$	4.234	1.134	0.405	9.942
$risk_{it}$	0.228	0.303	0.000	1.000
$cost_{ij}$	0.259	0.333	0.000	1.000
$funloc_{ij}$	0.069	–	0	1
$fungmt_{ij}$	0.259	–	0	1
$funeu_{ij}$	0.049	–	0	1
lp_{ij}	0.270	–	0	1
HHI_j	0.078	0.098	0.015	0.378
gp_{ij}	0.663	–	0	1
$rdpi_{ij}$	0.026	0.079	0.000	0.957

publicly available information. We denote the spillover variable by $spill_{ij}$. Unfortunately, the 2006 edition of the CIS does not collect any information on the level of appropriability, namely, the degree of strategic protection the firms adopt for their innovations. Table 3.3 reports summary statistics of the regressors used to estimate the set of coefficients θ and the variances $\mathbf{V_1}$ and $\mathbf{V_2}$ of the mixed logit model as in (3.5). The variable $spill_{ij}$ measures the degree of importance of publicly available source of information. The variable was originally coded from 0 (not used) to 3 (highly important), and we recoded in $\{0, 1/3, 2/3, 1\}$. Size turns out to be another important determinant of R&D cooperation. Cassiman and Veugelers (2006) highlighted how the size of the firms is a control variable, traditionally used by the literature in firm-level analysis. Therefore, in line with the existing literature, we include firm size measured as the logarithm of the number of the firm's employees ($size_{ij}$). On average the log of labor is 4.234. The innovativeness of firms is proxied by two dummy variables. The first dummy variable, rad_{ij}, proxying radical innovation, takes the value one if the enterprise introduced a new good or service into its market before its competitors, while the second dummy, inc_{ij}, is equal to 1 if a new good or service was already available from the competitors in the market of interest (incremental innovation). Firms introducing a radical innovations account for 32% of the sample, while imitative innovators sum up to 34% of the total number of firms in the sample. Among determinants of R&D cooperation, firms' absorptive capacity is considered as one of the most important. To proxy firms' absorptive capacity R&D personnel intensity ($rdpi_{ij}$) is often used (Tether, 2002; Belderbos et al., 2004b,a) instead of R&D expenditure. R&D personnel intensity is defined here as the ratio between R&D personnel and size of firms. On average, only 3% of the labor force is dedicated to research activities.

The variables $risk_{ij}$, and $cost_{ij}$ measure the degree of importance, $\{0, 1/3, 2/3, 1\}$, attributed by firms to the two factors hampering innovation activities. The cost factor seems to be larger than the risk associated with market uncertainty (26% versus 23%). The literature on R&D cooperation shows that the risks and costs of innovation and the need to exploit complementary resources are the main motivations for cooperative behavior, and therefore, that cooperative behavior may be positively related to a number of obstacles such as high risks and cost of innovation. R&D cooperations, in fact, allow firms to share costs or to reduce risks of innovation. In this regard, we hypothesize that a cooperation with customers could reduce the risk to introduce a radical innovation in the market. With regard to public funding, this in general has a positive influence on firms' R&D expenditure and, following Veugelers (1997), indirectly influences the propensity to cooperate in R&D. We therefore include dummies taking value 1 if the firm benefitted from both local, national, and European scientific and technological policies ($funloc_{ij}$, $fungmt_{ij}$, and $funeu_{ij}$, respectively) sponsoring cooperative projects, as they potentially constitute an incentive to cooperate. The number of firms receiving a national funding account for 26% of the observations, while local and European financial support is less frequent, as only 7 and 5 percent of firms received public sector support, respectively.

The integration of the firm into a group may also have a positive influence on cooperation as it indicates access to a substantial pool of resources, which are complementary to R&D. Thus, we include the intra-group variable, *gp*, a control variable which is equal to 1 when the firm belongs to a group, and 66% of the firms are part of an enterprise group.

As for sectoral characteristics, we include the degree of industry concentration, measured by the Herfindahl-Hirschman index (HHI_{ij}), as this may affect firms' motivation of combining resources with other firms. However, the impact of market concentration on the firms' propensity to form R&D alliances is a theme empirically less explored. The empirical contribution of Hernán et al. (2003) showed a positive impact of market concentration on the propensity of firms to cooperate in R&D, since a more concentrated industry offers a greater opportunity for internalization of spillovers. Wang and Zajac (2007), instead, did not achieve clear-cut conclusions since they found different results for different model specifications. The estimated Herfindahl-Hirschman index is, on average, 8%. The index range from, 0.015, not concentrated, to 0.378, moderately/highly concentrated sector.

3.5 Estimation Results

In the current section, we present the results of the multivariate mixed-effects logistic regression model as in (3.5), showing how R&D collective interactions are firm-level processes with high heterogeneity of actors and activities, where a strong sectoral specificity exists. In particular, we distinguish between firm-level and sectoral-level determinants of R&D cooperation. Table 3.4 reports results for the complete sample of 1929 observations.

As it has been shown in previous literature, some (if not all) of the variables included in this model may be endogenous. In this paper, we do not control for omitted variables, selection, or simultaneity, therefore the estimates are causally uninterpretable. The estimation approach we propose is rather aiming at the decomposition of the conditional variance structure due to firm- and sector-level heterogeneity. In fact, the expected correlation among the different firms' cooperative strategies is expected to have a nested structure that could pick the multiplier effect of innovation policies.

With the potential lack of interpretation in mind, we note that the estimated coefficients statistically differ across the equations[24], indicating the appropriateness of distinguishing between cooperation types. We observed that the innovativeness of a firm plays an important role in disentangling the determinants of R&D

[24]We formally test differences in the estimated coefficients using a Welch two-sample t-test.

cooperation. In particular, as expected, developing radical innovation, therefore being a 'true' innovator, has a larger impact on all four cooperation types than performing only imitative innovations. It is interesting to note that the enterprises which introduced a new good or service in their reference market before their competitors, tend to form alliances with customers in the first place. As a matter of fact, introducing a radical innovation increases the odds of cooperating with customers by more than 2.5 times (i.e., it increases the probability to cooperate with customers by almost 72%). On the other hand, introducing a new (to the firm) good already available from the competitors in the market in which the firm operates, enhances the chance to establish a formal R&D agreement with a competitor more than with other partners.

The hypothesis that incoming spillovers positively affect the probability of cooperation is confirmed only for two cooperation types. The spillover variable (*spill*), measured as the total pool of external knowledge that is potentially available for a firm, has a high and significantly positive impact only on competitors and institutional cooperation strategies (odd-ratios of 3.661 and 6.043, respectively). Higher incoming spillovers positively affect the probability of cooperating with research institutes and competitors, but have no effect on cooperation with customers or suppliers. Cassiman and Veugelers (2002) also do not find evidence of a significant impact of incoming spillovers on vertical cooperation, but do find statistical evidence of the positive relation between appropriability and the probability of cooperating with customers or suppliers. As a matter of fact, our results slightly differ from the existing studies in that, while they find a significant increase in the probability of cooperating with research institutes due to incoming spillovers, they do not find any significant effect on the propensity to collaborate with competitors. We suggest two different but correlated explanations to this phenomenon. The first concerns the construction of our spillover variable, the second relates to the multilevel structure of our model where firms are nested in sectors. The incoming spillovers are measured by the importance of publicly available information for the firm's innovation process, but consider as a source of information also the question relative to the importance of professional or industry associations for innovation activities. Such a construction of the spillover variable implicitly calls for a higher probability of horizontal cooperation. Indeed, a firm can exploit much better the information coming from an industry association in the context of a cooperative alliance within the same sector. On the other hand, we hypothesize that the coefficient of our spillover variable is not significant for vertical cooperation types in that, by definition, the variable does not consider inter-industry collaboration.

Given the multilevel structure of our model where firms are supposed to be nested in sectors, the spillover variable plays the most important role. Indeed, the synergy between this particular source of information and our innovative multilevel structure might be at the root of such a significant regression coefficient.

In line with empirical findings, firm size is positive and significant in all four cooperation strategies. Larger firms are more likely to have the right absorptive capacity required to engage in R&D cooperation, and this effect is stronger for cooperation with universities and suppliers. In line with previous studies we used the logarithm form, since it is natural that this effect is attenuated when the number of employees grows large.

Empirical literature generally found a positive impact of barriers to innovation such as costs and risks connected to the innovation process on the propensity of firms to cooperate. The propensity of firms in engaging in R&D cooperations with universities or other research institutions are expected to be positively correlated with the costs of innovation (Cassiman and Veugelers, 2002). Indeed, the cost sharing motive is found to be an important determinant for firms when deciding to cooperate with customers or research institutes. On the other hand, cooperations with customers, other than to access to complementary knowledge, are aimed to reduce the risk associated with bringing an innovation to the market. As a matter of fact, the risk-sharing variable is found to be significantly positive and the magnitude of the regression coefficient is much larger than the one of the cost-sharing variable. Further, the risk factor is not significant for the other three collaborative agreements.

Furthermore, as expected, R&D personnel intensity has a positive effect on the probability of cooperation with all sorts of partners.

The variables for public financial support for innovation activities have a positive and significant impact on almost all cooperation strategies. This may suggest that subsidies, especially those from the European Union, promote competitive

Table 3.4: Estimation results : Multivariate mixed logit

	Suppliers	Customers	Competitors	Institutional
Intercept	-3.541*** (-4.448, -2.568)	-4.691*** (-5.721, -3.637)	-5.718*** (-7.015, -4.441)	-6.085*** (-7.365, -5.045)
rad	0.593*** (0.307, 0.898)	0.926*** (0.621, 1.220)	0.618*** (0.221, 1.054)	0.593*** (0.255, 0.899)
inc	0.280** (0.014, 0.577)	0.486*** (0.172, 0.809)	0.596** (0.196, 1.043)	0.410*** (0.103, 0.756)
size	0.339*** (0.222, 0.461)	0.290*** (0.161, 0.428)	0.225*** (0.068, 0.379)	0.343*** (0.212, 0.489)
spill	0.416 (-0.145, 0.882)	0.539 (-0.053, 1.193)	1.298*** (0.321, 2.214)	1.799*** (1.021, 2.644)
cost	0.197 (-0.118, 0.521)	0.342* (-0.036, 0.664)	-0.145 (-0.594, 0.267)	0.630*** (0.243, 1.002)
risk	-0.098 (-0.397, 0.226)	0.581*** (0.243, 1.001)	0.196 (-0.240, 0.704)	-0.072 (-0.393, 0.289)
rdpi	2.184*** (0.908, 3.562)	0.713 (-0.237, 1.648)	0.646 (-0.540, 1.660)	2.165** (1.032, 3.464)
fungmt	0.532*** (0.254, 0.857)	0.565*** (0.235, 0.891)	0.353* (-0.009, 0.750)	0.724*** (0.384, 1.030)
funeu	0.641*** (0.153, 1.188)	0.301 (-0.218, 0.838)	1.133*** (0.622, 1.646)	0.804*** (0.332, 1.353)
gp	0.411*** (0.106, 0.758)	0.1189 (-0.165, 0.541)	0.141 (-0.281, 0.601)	0.494*** (0.125, 0.853)
		Firm-level Random Intercept		
$\hat{\sigma}_{\gamma}$	7.532*** (6.797, 8.297)	5.324*** (4.519, 6.750)	2.970*** (2.443, 3.457)	3.671*** (3.362, 4.615)
		Sector-level Random effects (Averages over sectors)		
$\hat{\sigma}_{\alpha_0}$	0.634*** (0.273, 1.665)	0.701*** (0.279, 1.817)	0.676*** (0.245, 1.729)	0.614*** (0.272, 1.633)
$\hat{\sigma}_{lp}$	0.326*** (0.210, 0.842)	0.329*** (0.200, 0.855)	0.382*** (0.188, 0.909)	0.324*** (0.200, 0.836)
$\hat{\sigma}_{HHI}$	0.638*** (0.257, 1.666)	0.709*** (0.301, 1.917)	0.684*** (0.225, 1.661)	0.644*** (0.279, 1.761)

Note: 95% Posterior Credible Interval (PCI) in brackets; significance codes for the PCIs : 0.01 '***'; 0.05 '**'; 0.1 '*'; DIC = 5516.991

R&D partnerships, in particular with research institutions. Lastly, we find that firms belonging to a domestic group are more likely to cooperate with suppliers or research institutes.

3.5.1 Firm- and Sector-level Heterogeneity

The hypothesis of a more complex structure of the heterogeneity of cooperation determinants is confirmed by our results. Both firm- and sector-level variances and covariances are found to be significant, meaning that enterprises within the same industry share similar characteristics (same random effects), which lead to correlation between research partners' choices.

The proportion of variation explained by the firm- and sector-level random intercepts varies across cooperation strategies. The firm- and sector-level variations of cooperating with suppliers account for the 86.5% and the 2.8%, respectively, 82.0% and 6.6% for cooperating with customers, 71.2% and 3.9% for horizontal research alliances, and 75.5% and 3.1% are the firm- and sector-level proportions of residual variance specific of cooperating with a public research institution.

In general, we find that the variance of the firm-level intercept is much larger than that of the sector-level. Concerning the coefficient estimates of sector-level variables, namely the Herfindhal index and the measure of legal protection, are reported at the bottom of Table 3.4. Both the legal protection measures adopted by the firm and the degree of industry concentration affect their motivation of combining resources with other firms.

In figure 3.1 we plot the posterior distributions of the sector-level correlations between the 4 different cooperation strategies. The plots on the left show time series of the values of 1600 samples of the posterior distribution correlations. These graphs constitute a powerful visual inspection tool. Indeed, we can examine these MCMC draws to check that `MCMCglmm`'s algorithms fitted the data and our model specification quite well. Indeed, the posterior distributions seem to be mixing well around the mean values of the correlation coefficients. The plots on the right show the same data on the correlation parameter draws as marginal posterior distributions (kernel density estimates).

Table 3.5: Firm- and Sector-level random effects correlation matrix

Coop. type	Suppliers	Customers	Competitors	Institutional
Suppliers	1.000	0.962*** (0.908, 0.988)	0.987*** (0.966, 0.999)	0.889* (0.839, 0.948)
Customers		1.000	0.926*** (0.882, 0.966)	0.740*** (0.611, 0.860)
Competitors			1.000	0.931*** (0.874, 0.987)
Coop. type	Suppliers	Customers	Competitors	Institutional
Suppliers	1.000	0.723*** (0.346, 0.974)	0.724*** (0.306, 0.965)	0.714* (0.310, 0.973)
Customers		1.000	0.708*** (0.286, 0.975)	0.741*** (0.372, 0.979)
Competitors			1.000	0.690*** (0.274, 0.963)

Note: 95% Posterior Credible Interval (PCI) in brackets; significance codes for the PCIs : 0.01 '***'; 0.05 '**'; 0.1 '*'.

The values of correlations between the different cooperation strategies at the both firm- and sector-level are summarized in Table 3.5. The upper part of the table reports the correlation between cooperation choices at the level of the enterprise. The firm-specific correlation coefficients are on average 30% higher the sector-specific correlation coefficients.

Additionally, comparing the estimates of the multivariate mixed logit with the multivariate probit (Table 3.7), we notice that, in general, for all cooperation strategies, developing radical innovation or performing imitative innovations do not have the the same strong effect on the propensity to cooperate as with the multivariate mixed logit estimates. Regression coefficients' estimates are statistically significant at the 5% only for cooperations with customers or suppliers. Another distinguishing feature of the multivariate probit consists in the fact that industry legal protection does not seem to play any remarkable role in enhancing the chance to cooperate with suppliers, competitors or research institutes. On the other hand, the level of industry concentration (HHI) appears to have a positive and significant impact on each of the R&D alliances, except for the one with competitors. This result is counterintuitive as we would expect a significant and high effect of the market concentration especially on the probability to cooperate horizontally with competitors, as such an alliance would offer a greater scope for internalization of spillovers (Hernán et al., 2003).

3.6 Evaluating the impact of public funding on innovative output

The success of all stages of innovation should be perceived as the outcome of a collaborative occurrence, arising from cognitive proximity (Cohen and Levinthal, 1990). Proximity can be defined, other than from a geographical perspective, in terms of organizational and institutional proximity. In particular, (Dosi, 1999) claims that the production of innovative output is influenced by the "social embeddedness" of firms' routines and strategies. Simply put, the innovativeness of a firm is likely to be driven by socially specific factors, such as the nature of the local labor markets, workforce training institutions, financial institutions, mechanisms governing the support of business start-ups and development, science and technology policies, inter-firm and firminstitutional interactions.

Large part of European policy measures supporting innovation activities focus on both the promotion of R&D cooperation between the actors of the innovation system (which includes enterprises, universities, and research institutes), and on the improvement of the conditions for the uptake of innovations and/or to improve the articulation of demand in order to spur innovations and the diffusion of innovations.

Policy evaluation of European initiatives, regional, or national innovation policies usually relies on counterfactual analysis, and on the implementation of econometric techniques such as matching (Heckman et al., 1998; Klette et al., 2000; Almus and Czarnitzki, 2003). Undertaking such a systematic approach goes beyond the scope of this paper. We rather want to show how the proposed hierarchical structure has an impact on the measured effects innovation policies on both cooperation propensity and innovative output.

In particular, in this paragraph, we investigate the relationship between innovative output, measured as the percentage of firm's total turnover from goods and service innovations introduced during 2004 to 2006 that were new to the firm's market ($Inno$), and innovative input, simply measured by the R&D personnel intensity ($rdpi$). We then control for public financial support for innovation activities at different levels of government (regional, $funloc$; central government, $fungmt$; European Union, $funeu$) and for cooperation with four types of partners. Therefore, the linear mixed model describing the impact of public funding on innovative sales is:

$$Inno_{ij} = \beta_0 + \beta_c' y_{ij} + \beta_r rdpi_{ij} + \alpha_j' F_{ij} + \nu_{ij}, \tag{3.6}$$

where $y_{ij} = (y_{ij}^1, \ldots, y_{ij}^4)'$ is the vector of 4 binary variables taking value 1 when one of the 4 cooperation partner is selected. As before, we define α_j as a normally distributed sector-level random effect, and $F_{ij} = (funloc_{ij}, fungmt_{ij}, funeu_{ij})'$ as the vector of dummies proxing whether the firm received a funding from a regional, governmental, or European institution. The error term ν_{ij} is assumed to be normally distributed.

Table 3.6 reports the estimates of two models. The first column displays estimated mean values and 95% posterior intervals of a linear mixed model without sector-level random components[25]. The second column presents estimates of a linear mixed model with random effects and random intercept. Specifically, the effect of variables proxying for regional, central government and European Union R&D support ($funloc$, $fungmt$, $funeu$) is considered to be random at the sectoral level. We also allow for the sector-specific intercept to be random.

Results confirm our beliefs. When considering a hierarchical framework, the policy interventions are relevant for innovative output. The realizations of sectoral intercept and slopes of the policy coefficients are expected to lie in 95% positive intervals. That is, the marginal effect of a regional innovation policy on the percentage of innovative turnover will be between 0.032 and 0.072. The same way of reasoning applies for the other policies.

[25]Since the MLE estimator and the mean of the posterior are asymptotically equivalent and their difference depends on the inverse of the square root of the sample size, the larger the sample size the narrower this difference. As our sample is pretty large (1929 observations), this difference is likely to be negligible

Table 3.6: Estimation results : Mixed linear model

Dep. var.:Y_{Inno}	*Fixed Coefficients*	*Random Coefficients*
Co^{supp}	-0.011 (-0.030, 0.012)	-0.008 (-0.029, 0.012)
Co^{cust}	0.030*** (0.008, 0.051)	0.025** (0.004, 0.047)
Co^{comp}	-0.007 (-0.030, 0.020)	-0.007 (-0.032, 0.015)
Co^{inst}	0.022** (0.002, 0.043)	0.027*** (0.006, 0.048)
rdpi	0.007*** (0.002, 0.013)	0.008*** (0.003, 0.014)
funloc	0.006 (-0.017, 0.028)	0.050***($\widehat{\sigma}_{loc}$) (0.032, 0.072)
fungmt	0.020*** (0.003, 0.035)	0.045***($\widehat{\sigma}_{gmt}$) (0.029, 0.062)
funeu	0.007 (-0.018, 0.033)	0.054***($\widehat{\sigma}_{eu}$) (0.034, 0.078)
Intercept	0.045*** (0.034, 0.057)	0.070***($\widehat{\sigma}_{int}$) (0.039, 0.107)
DIC	-1480.567	-1477.612

Note: 95% Posterior Credible Interval (PCI) in brackets; significance codes for the PCIs : 0.01 '***'; 0.05 '**'; 0.1 '*'

It is also interesting to note that, when disregarding the sector-specific random effects(first column), the central government R&D policy resulted to have the higher (and statistically significant at 1% level) marginal effect on innovative output (0.020, compared with 0.006 for regional funding, and 0.007 for European funding). However, when turning to our multilevel approach that can take into account organizational proximities within sectors, the impact of European Union fundings has the biggest HPD interval (0.034, 0.078). European scientific and technological policies increase the innovative sales by a value falling in the aforementioned interval.

3.7 Conclusions

Using data from the last available 2006 edition of the Community Innovation Survey for the Netherlands, this paper contributes to the existing empirical literature, by proposing a methodology to study the determinants of innovative collaborative agreements and to assess the impact of public financial support to R&D. In particular, we explore the firm- and the sector- level heterogeneity of the determinants of either forming an R&D alliance, or selling innovative products, by considering a (generalized) linear mixed model.

The two steps of our analysis can be summarized as follows. In the first stage, we investigate the relationship between the factors driving the propensity to collaborate with different research partners, assuming a multivariate hierarchical logit model. The second step confirms the key role of the assumed multilevel structure, by considering the relationship between innovative output and innovative input, controlling for public financial support to innovation activities at different levels of government.

To our knowledge, this is the first attempt to model the both firm- and sector-level heterogeneity in the determinants of R&D partner's choices and innovation output.

Our hypothesis of a heterogeneity across firms and sectors is confirmed by the results. All covariances are found to be significant. In other words, firms within the same industry share similar characteristics (same random effects), which lead to correlation between research partners' choices.

This confirms that R&D cooperation, as well as the innovative production, is a firm-level process, where a strong sectoral specificity exists.

Taking into account this sectoral-organizational proximity can help assessing the right impact of R&D policies on innovative output. We have seen that, when using the multilevel approach the impact of public fundings has a positive and significant sign, while, when omitting this nested framework the policies have a poor effect on innovative turnover.

Moreover, our suggested empirical framework can be brought to a deeper level of analysis, if data on markets were observed. As a matter of fact, understanding market dynamics could be the key to create more innovation-friendly market conditions that are necessary to reduce the time-to-market of new goods and to enable emerging sectors and/or markets to grow faster.

In these markets, for example, the removal of barriers would essentially contribute to the competitive process and lead to the emergence of new markets. Competitiveness is here meant not only as the ability of the firm to come up with innovation from its internal technological strength, but also on its ability to access the innovation networking, that, as we have shown, depends on sector-specific networking and proximity (and, presumably, also on market-specific characteristics).

3.8 Appendix

3.8.1 Variable Description

- *Cooperation*: Dummy variable which takes value 1 if the firm has declared to cooperate with at least one of the possible research partners, i.e., suppliers, clients, competitors, consultors, universities, research institutes.

- *Cooperation with competitors*: Dummy variable which takes the value 1 when the firm has actively participated with its competitors on innovation activities.

- *Cooperation with customers*: Dummy variable which takes the value 1 when the firm has actively participated with its clients or customers on innovation activities.

- *Cooperation with suppliers*: Dummy variable which takes the value 1 when the firm has actively participated with its suppliers of equipment, materials, components, or software on innovation activities.

- *Cooperation with research institutions*: Dummy variable which takes the value 1 when the firm has actively participated with universities or other higher education institutions, or government or public research institutes on innovation activities.

- *Incoming Spillovers*: Variable which takes the value 0 if innovation ideas are not originated by Professional conferences, exhibitions, meetings and journals, professional and industry associations. The variable was originally coded from 0 (not used) to 3 (highly important), and we recoded in 0, 1/3, 2/3, 1.

- *Legal Protection*: Dummy variable that takes value 1 if the firm applied for a patent, registered an industrial design, or a trademark, or claimed a copyright.

- *Size*: Log of number of employees of the firm.

- *Cost*: Variable measuring the importance of the costs of innovation, or the lack of funds, or access to finance in hampering the firm's innovation activities or projects or influencing the decision not to innovate. The original variable takes values between 1 (high) and 4 (not relevant). Rescaled between 0 (not relevant) and 1 (high).

- *Risk*: Variable measuring the importance of the uncertainty of the demand for innovative goods or services in hampering the firm's innovation activities or projects or influencing the decision not to innovate. The original variable takes values between 1 (high) and 4 (not relevant). Rescaled between 0 (not relevant) and 1 (high).

- *Radical innovator*: Dummy variable that takes value 1 if the enterprise introduced a new or significantly improved good or service onto its reference market before the competitors.

- *Incremental innovator*: Dummy variable that takes value 1 if the enterprise introduced a new or significantly improved good or service that was already available from the competitors in its reference market.

- *Absorptive capacity*: It is proxied by R&D personnel intensity, measured as the ratio between the log of researchers (full time equivalent) and the size of the firm.

- *Enterprise Group*: Dummy variable that takes value 1 if the firm is part of an enterprise group. A group consists of two or more legally defined enterprises under common ownership. Each enterprise in the group may serve different markets, as with national or regional subsidiaries, or serve different product markets. The head office is also part of an enterprise group.

- *Public funding*

  - *Regional funding*: Dummy variable taking value 1 if the firm received any public financial support for innovation activities from local or regional authorities;

  - *National funding*: Dummy variable taking value 1 if the firm received any public financial support for innovation activities from central government (including central government agencies or ministries);
  - *European funding*: Dummy variable taking value 1 if the firm received any public financial support for innovation activities from the European Union.

- *Industry concentration*: As a measure of industry concentration we adopt the Herfindahl-Hirschman Index, computed as the sum of the squared firms' market shares.

- *Innovative output*: Innovation output is proxied by the percentage of total turnover from product or process innovation (new to the firm and/or to the market)

3.8.2 Bayesian Multivariate Mixed Logit Model

The maximum likelihood method is the standard approach for statistical inference in the mixed effects model. In order to maximize the sample likelihood, integration over the random-effects distribution must be performed. Yet, there exists no analytical solution for the intractable integral in equation (3.5). As a result, estimation is much more complicated than in models for continuous normally distributed outcomes where the solution can be expressed in closed form. Various approximations for evaluating the integral over the random-effects distribution have been proposed in the literature; many of these are reviewed in Rodríguez and Goldman (1995).

Simulation methods are also popular techniques to estimate mixed effects models (Train, 2009). The unconditional probabilities in equation (3.5) are approximated through simulation for any given value $\boldsymbol{\theta}$ of the parameters of the mixing distribution $f(\boldsymbol{\alpha}|\boldsymbol{\theta})$. Such methods fall under the rubric of Markov Chain Monte Carlo (MCMC) algorithms.

In this paper we adopt a Bayesian approach and explore the MCMC fitting of the multivariate mixed logit model. One advantage of the Bayesian approach over its frequentist counterpart includes the fact that the Bayesian procedures do not require maximization of any function. For complicated random effects structures, computation of a single maximum likelihood fit can be expensive, making the simulation of statistics of interest computationally prohibitive. Second, with Bayesian procedures, estimation properties, such as consistency and efficiency, can be attained under more relaxed conditions than with classical procedures. As shown in Train (2009) (Chapter 10), consistency of the Maximum Simulated Likelihood (MSL) estimator depends on the relationship between the number of draws that are used in the simulation and the sample size. If the number of draws is considered fixed, then the MSL estimator does not converge to the true parameters, because of the simulation bias. The simulation bias disappears as the sample size rises without bound together with the number of draws. In contrast, the Bayesian estimators are consistent for a fixed number of draws used in simulation and are efficient if the number of draws rises at any rate with sample size.

Following the Bayesian approach, the model parameters $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, $\mathbf{G}$, summarized in the vector $\boldsymbol{\theta}$, are treated as random variables. The assumed distributions for the parameters, called *prior distributions* and denoted by $f(\boldsymbol{\theta})$, borrow information from past studies, logic, or from the researcher's ideas about the values of these parameters. Therefore, the prior distribution represents how likely the researcher thinks it is for the parameters to take a particular value, over all possible values that the parameters can take. Bayesian inference is based on the *posterior distribution*, $f(\boldsymbol{\theta}|\mathbf{y})$, which is the conditional distribution of the conjectured, but unknown, parameters $\boldsymbol{\theta}$, given the observed data $\mathbf{y} = y_1, \ldots, y_n$.

The choice of a prior distribution $f(\boldsymbol{\theta})$ affects Bayesian estimation. In other words, Bayesian inference may be influenced by a "strong" prior. In absence of any prior information, a non-informative prior is chosen ($f(\boldsymbol{\theta}) \propto 1$)) and Bayesian inference is asymptotically equivalent to likelihood inference. In practice, we always specify a diffuse prior for $\boldsymbol{\beta}$, and try different values of the set of parameters $\boldsymbol{\alpha}$, $\mathbf{G}$, as a sensitivity analysis.

To estimate the parameters of the Generalized Linear Mixed Model (GLMM) defined in Section 4 following a Bayesian approach (Zeger and Karim, 1991; Gelman et al., 2003), we use the R package `MCMCglmm` (Hadfield and Kruuk, 2010). The default prior chosen by `MCMCglmm` for the regression model parameters β^c is a non-informative, normal distribution $N(0, 1e+10)$[26], while for both the residual and random-effect variance matrices a diffuse inverse-Wishart distribution is assumed, which is commonly used in practice. Then, assuming that the priors are independent,

$$f(\boldsymbol{\beta}, G) = f(\boldsymbol{\beta})f(G), \tag{3.7}$$

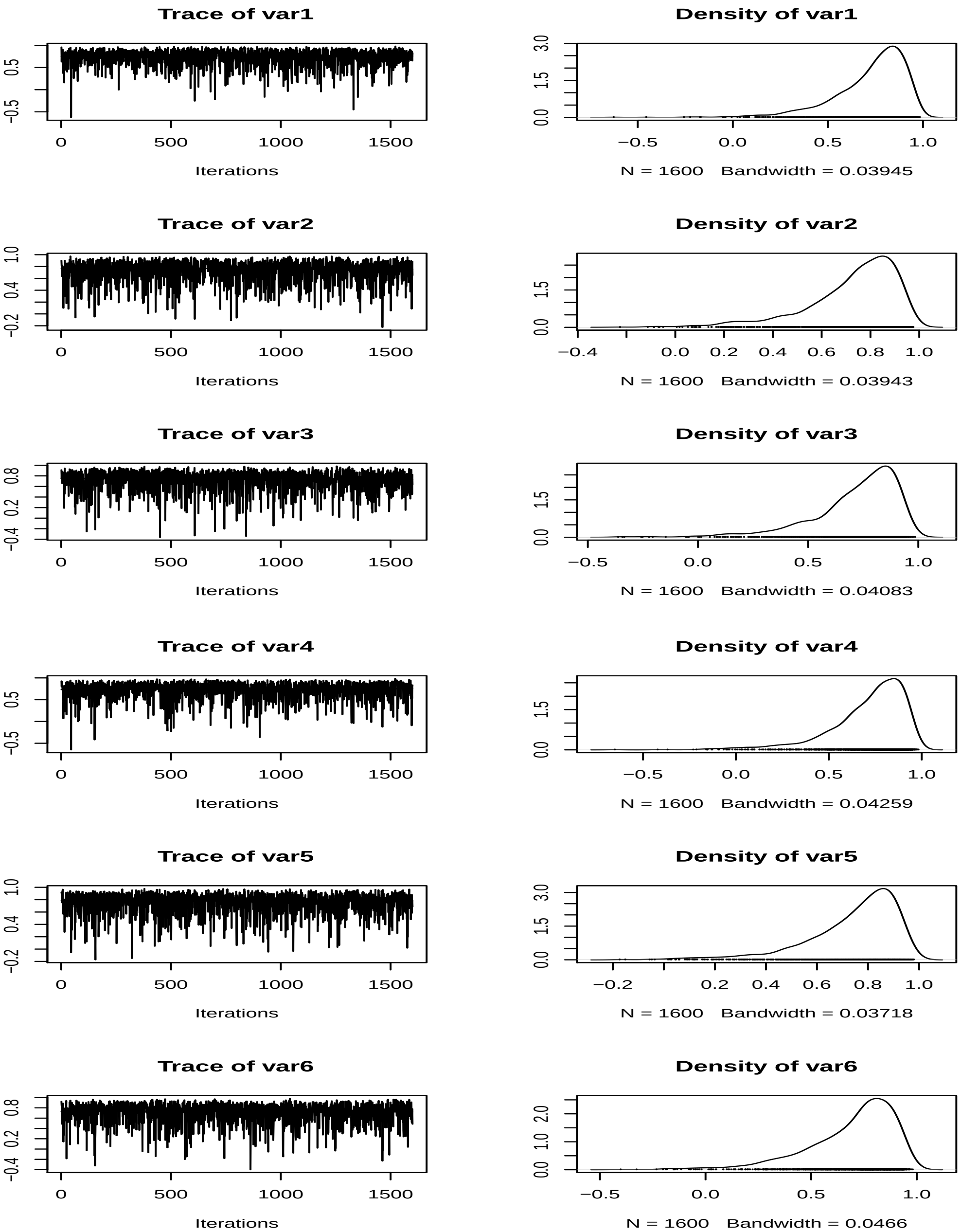
the posterior distribution can be written as

$$f(\boldsymbol{\beta}, G, \boldsymbol{\alpha}|\mathbf{y}) \propto \prod_{j=1}^{n}\prod_{c=1}^{4}\prod_{i=1}^{n_j^c} f_{ij}^c(y_{ij}^c|\boldsymbol{\alpha}_j^c, \boldsymbol{\beta}^c)f(\boldsymbol{\beta}^c)\prod_{j=1}^{n}\prod_{c=1}^{4} f(\boldsymbol{\alpha}_j^c|\mathbf{G})f(\mathbf{G}). \tag{3.8}$$

The R package `MCMCglmm` generate samples from the posterior distribution using Metropolis-Hastings updates (for more details on the sampling schemes, see Hadfield and Kruuk (2010)). Beginning with the starting values $(\boldsymbol{\beta}^{(0)}, \boldsymbol{\alpha}^{(0)}, W^{(0)})$, after a warm-up (also called "burn-in") period, we store a sample of $(\boldsymbol{\beta}, \boldsymbol{\alpha}, W)$ from the posterior distribution. Once we generate a large number of samples, the posterior mean and posterior covariance can be approximated by the sample mean and the sample covariance based on the simulated samples. Convergence of the MCMC sampling scheme was assessed using empirical and test-based approaches (Heidelberger and Welch, 1983; Geweke, 1992). Results from convergence diagnostics indicated that it was sufficient to burn-in the first 15,000 samples and take the subsequent 1,600 samples for inference.

[26]One of the many advantages of the package `MCMCglmm` resides in the great flexibility in the specification of various residual and random-effect variance structures. `MCMCglmm` allows variance structures of the form $\mathbf{G}=\mathbf{V}\otimes\mathbf{A}$: unstructured and completely parameterized covariance matrices. However, binary responses pose a special problem because the residual variance cannot be estimated because the variance is uniquely determined by the mean. Therefore, following Hadfield and Kruuk (2010), we apply restrictions on the prior distribution of the residual covariance matrix. In particular, we fix the parameters of the prior distribution at some value (1 for variances and 0 for covariances).

Figure 3.1: Correlation Estimates



Note: Var1: correlation between Suppliers and Customers; Var2: correlation between Suppliers and Competitors; Var3: correlation between Suppliers and Research Institutes; Var4: correlation between Customers and Competitors; Var5: correlation between Customers and Research Institutes; Var6: correlation between Competitors and Research Institutes.

Table 3.7: Estimation results : Multivariate probit

Variable	Customers		Suppliers		Competitors		Institutional	
rad	0.351***	(0.072)	0.243***	(0.074)	0.198*	(0.108)	0.161	(0.104)
inc	0.242***	(0.054)	0.090	(0.061)	0.170*	(0.091)	0.138*	(0.073)
size	0.110**	(0.043)	0.160***	(0.023)	0.143***	(0.050)	0.190***	(0.028)
spill	0.389**	(0.183)	0.474***	(0.076)	0.482*	(0.264)	0.942***	(0.198)
cost	0.153**	(0.076)	0.157*	(0.080)	-0.083	(0.121)	0.293**	(0.115)
risk	0.259***	(0.072)	-0.135***	(0.052)	0.010	(0.099)	-0.053	(0.127)
rdpi	0.173	(0.233)	0.574***	(0.213)	0.273	(0.295)	0.744***	(0.163)
fungmt	0.168**	(0.067)	0.152**	(0.074)	0.085	(0.112)	0.236**	(0.095)
funeu	0.199	(0.151)	0.420***	(0.079)	0.557***	(0.143)	0.438***	(0.113)
gp	0.114	(0.094)	0.208**	(0.097)	0.044	(0.140)	0.277***	(0.105)
lp	0.151***	(0.051)	0.090	(0.055)	0.047	(0.069)	0.141*	(0.074)
HHI	0.984***	(0.346)	0.449**	(0.183)	0.341	(0.366)	0.792***	(0.232)
Intercept	-2.276***	(0.158)	-1.866***	(0.155)	-2.601***	(0.193)	-3.086***	(0.243)
				Correlations Matrix				
	Customers		Suppliers		Competitors		Institutional	
Customers	1		0.761 ***	(0.041)	0.637 ***	(0.021)	0.640 ***	(0.037)
Suppliers			1		0.652 ***	(0.040)	0.686 ***	(0.027)
Competitors					1		0.615 ***	(0.051)

Note: Standard errors in parentheses; sample of 1354 firms; Log pseudolikelihood = -1959.293; Wald $\chi^2(6)$=824.763, Prob> χ^2 = 0.000.

Chapter 4

Identifying Firm-level R&D Cooperation and Innovation Decisions

Abstract. This paper investigate the distribution of the fixed and sunk costs associated with the firm–level decisions of innovating, spending, and cooperating in R&D, adopting a dynamic structural framework. The basic idea of the paper is to model the firms' decisions to cooperate in R&D and to innovate with a dynamic discrete choice model. None of the existing studies on heterogeneity of cooperation strategies or innovation processes deals with the nontrivial dynamics deriving from uncertainty and sunk costs of investments. Identifying the firms' primitives on productivity and investment decisions is key to have an encompassing understanding on what are the determinants on which the firm bases its choice to innovate and/or cooperate. Additionally, the suggested structural framework of firm heterogeneity in cost functions offers a straightforward extension to policy impact evaluation.

4.1 Introduction

An important source of productivity differentials across firms is related to R&D and innovation activities. Many authors have studied the connection between spending for R&D and productivity growth (Griliches, 1980; Jones and Williams, 1998; Hall and Mairesse, 1995; Crépon et al., 1998). As a result, a large number of empirical studies estimates the effect of R&D investment on such growth, finding that R&D spending has a significantly positive effect on productivity growth, with a rate of return that is about the same size as (or to some extent larger than) the rate of return on conventional investments. Crépon et al. (1998), examining the structural links between productivity, innovation input, and innovation output at the firm level, find that the firm innovation output rises with its research efforts, and the firm productivity correlates positively with a higher innovative output.

Nonetheless, before Ericson and Pakes (1995), most of the empirical literature in industry dynamics assumes that firms are endowed with an exogenous level of productivity. The "lucky" firms with high productivity survive and prosper, the others fail, and eventually exit the market.

The modern IO literature relaxes this exogeneity assumption by letting the productivity to be dependent on the investment decisions, so as to enhance the firms' survival chance (Ericson and Pakes, 1995; Doraszelski and Jaumandreu, 2008; Aw et al., 2011). Typically, in this context, the investment taken into consideration is past R&D expenditure (Doraszelski and Jaumandreu, 2008), or both R&D expenditure and export market participation (Aw et al., 2011). However, the firm that wants to survive must not only be innovative, but also ready to outsource knowledge and develop research networks. In fact, firms increasingly rely on the external acquisition of new technological knowledge, as the institutional locations of such resources can be quite disparate. Although not the primary source of produced knowledge, R&D outsourcing[27] (or external R&D)

[27] R&D outsourcing refers to the contractually agreed, non-gratuitous and temporary performance of R&D tasks for a client primarily by private contract research and technology organizations, but also by some private non-profit and related hybrid organizations (Howells, 1999; Grimpe and Kaiser, 2010)

has considerably increased in importance and accounts for a substantial share of the total innovation expenditure in a large number of firms.[28]

Therefore, in this paper we construct a model where firms invest in R&D activities with or without a research partner to improve their productivity levels. In particular, we develop and estimate a structural dynamic monopoly model to quantify the linkages between R&D spending, innovation and cooperation investment choices, and endogenous productivity. To our knowledge, our paper constitutes the first attempt to explicitly model the different collaborative R&D investment decisions adopting a dynamic structural framework.

All the other empirical studies aimed at determining whether research collaborations affect firm-level productivity rely on reduced-form regression approach. For example, Belderbos et al. (2004b), using data from two waves of the Dutch Community Innovation Survey (1996, 1998), analyze the impact of R&D cooperation on firm performance, regressing two measures of firm-level productivity growth on four different cooperation strategies. They also control for the effect of both own R&D efforts as well as the impact of incoming knowledge flows that are not due to cooperation. Carboni (2012) explores the variables that determine a firms R&D collaborative expenditure, in a regression analysis framework, correcting for heteroscedasticity and non-normality when dealing with a large number of zero response data.

Differently from these studies, the model we propose derives the firms' optimal R&D investment decisions where these depend on the past R&D activities and on the past level of productivity. Additionally, within the suggested framework we are able to model and retrieve the current fixed or sunk costs relative to the different (collaborative) R&D activities.

The literature on R&D cooperation shows that the risks and costs of innovation and the need to exploit complementary resources are the main motives for cooperative behavior, and therefore, that cooperative behavior may be positively related to a number of obstacles such as high risks and cost of innovation (Amoroso, 2011; Belderbos et al., 2004a,b). R&D cooperations, in fact, allow

[28]Source: Eurostat. "Innovation in Europe. Results for the EU, Iceland and Norway."

firms to share costs or to reduce risks of innovation. In this regard, we hypothesize that cooperating in research could reduce both the fixed and the sunk costs of introducing an innovation in the market.

We merge data on sales and factor inputs of Dutch manufacturing firms extracted from the Production Survey (PS), and three waves of the Community Innovation Survey (CIS) for the Netherlands, covering the period from 2002 to 2008. The leading sectors (chemicals, agri-food, transport, high-tech) in the Dutch manufacturing industry heavily depend on research and innovation, and these are, in turn, driven by a wide range of factors, such as firm performance, market conditions, policy interventions, and government requirements to reduce environmental damages. In this paper, we assume the firm bases its decision of engaging in R&D or in innovation with or without a research partner on past choices, firm-level total factor productivity, and a demand shifter, proxying for the industry characteristics.

In Section 2 we present the model that we use to retrieve information on both fixed and sunk costs, and consequently on the optimal R&D, innovation, and cooperation decisions. In Section 3, we discuss the empirical strategy used to retrieve estimates of the static parameters of the model. Namely, we illustrate how we obtain a measure of firm-level productivity, demand elasticity, and an aggregate demand shifter. Moreover, we present estimates of the fixed costs associated with each investment choice, in the static case, i.e., when the firm does not take into account the future payoffs in its profits maximization. Section 4 describes the steps of the algorithm developed by Imai et al. (2009) that is used to obtain the dynamic parameters estimates. Section 4 and 5 describe the data and the results, respectively. In Section 6 we present the results for a policy simulation and the last section concludes.

4.2 Structural Framework

The empirical model builds on the class of models of dynamic entry games in IO, where the dependent variable is the firm's decision to enter or not in a market. In the same spirit, this paper defines the entry decision as the adoption of a set of discrete decisions: investing in research and development (R&D), cooperating, innovating, and innovation cooperation. These decisions are assumed to be costly to reverse and, therefore, associated with sunk costs. As firms are assumed to be forward looking, they take into account the implications of their decisions (and the associated costs) on their future payoffs. Time is discrete and indexed by t. The single-agent dynamic optimization problem is solved for the N firms operating in the market, which we index by $i \in I = \{1, 2, \ldots, N\}$. Following the standard setting of Ericson and Pakes (1995), and adapting it to a monopolistic competitive setting, firms compete on two different dimensions: a static and a dynamic dimension. Within the dynamic dimension, a firm makes the investment choice indexed by $k \in \{na, rd, c, d, cd\}$, where the vector of choices is defined as $a_{it} = (na_{it}, rd_{it}, c_{it}, d_{it}, cd_{it})'$, with $a_{it} \in A_i \equiv \{0, 1\}^5$. The firm-specific choice na_{it} takes value one if the firm does not engage in any activity other than operating in the market; rd_{it} takes value one if the firm decides to spend in R&D; choices c_{it} and d_{it} match firms' decisions to start a research collaboration and to invest in a technological upgrade, respectively; action cd_{it} tags the decision to both innovate and cooperate (with either another firm or a research institute or a supplier/customer).

4.2.1 Static decisions

In every period, firms are competing in prices in a static Bertrand model. Let P_{it}, the price, be the static decision variable of firm i at time t. The demand curve faced by the monopolistically competitive firm is assumed to follow a Dixit–Stiglitz form:

$$Q_{it}^D = Q_t^j (P_{it}/P_t^j)^\eta e^{u_{it}^d} \tag{4.1}$$

where Q_{it}^D is the demanded quantity for a firm i, Q_t^j and P_t^j are the sector j aggregate production and price index, respectively, $\eta < -1$ is the constant elasticity of demand, and u_{it}^d is a demand shock.

The production function is assumed to take the form of a Cobb-Douglas, with the gross output Q_{it} of firm i at time t function of three specific inputs and productivity:

$$Q_{it} = A_{it} K_{it}^{\theta_{iKt}} L_{it}^{\theta_{iLt}} M_{it}^{\theta_{iMt}}, \tag{4.2}$$

where K_{it} denotes capital, L_{it} labor, and M_{it} intermediate goods, consisting of materials and energy, for firm i at period t. $\theta_{iKt}, \theta_{iLt}, \theta_{iMt}$ are the elasticities of output with respect to capital, labor, and intermediate goods, respectively. A_{it} represents the Hicksian neutral efficiency level of firm i at time t. The logarithm of A_{it} is defined as $A_{it} \equiv \exp(\theta_0 + \omega_{it})$ and is defined as the sum of the mean productivity level across firms and over time, θ_0, and the productivity shock which is observable by the firm, but not to the econometrician (for example, managerial ability, quality of research), ω_{it}..

Following the literature on imperfect competition in both product and labor markets (Bughin, 1993, 1996; Crépon et al., 2002; Dobbelaere, 2004; Abraham et al., 2009; Dobbelaere and Mairesse, 2011; Amoroso et al., 2012), we relax the conventional assumption of perfect competition in the labor market, allowing both firms and workers' union to have some market power. The workers bargain with the firm over both the levels of employment, L_{it}, and of the wage, W_{it}. Additionally, we define the firm level profits as

$$\Pi_{it} \equiv P_{it}Q_{it} - W_{it}L_{it} - FC(K_{it}, M_{it}, a_{it}), \tag{4.3}$$

where $FC(\cdot)$ are the (avoidable) fixed costs (costs that do not vary with the quantity of output produced, but are not irrevocably committed; (Wang and Yang, 2001)), depending on capital, material, and innovative investment. Moreover, we define the union's utility function as

$$U_{it}(W_{it}, L_{it}) \equiv L_{it}(W_{it} - \bar{W}_{it}),$$

where $\bar{W}_{it}$ is the reservation wage. Finally, the efficient bargaining model can be written as a weighted average of the logarithms of workers' aggregate gain from union membership and the firm's profits:

$$\max_{L_{it}, W_{it}} \left[\phi_{it} \log(U_{it}(W_{it}, L_{it})) + (1 - \phi_{it}) \log \Pi_{it}\right],$$

where $\phi_{it} \in [0, 1]$ is the degree of union bargaining power. In the static setting, the firm maximizes only with respect to the variable costs, namely, the cost of labor. Amoroso et al. (2012) show that, maximizing with respect to labor, and taking into account the demand curve faced by the monopolistically competitive firm, results in the following expression for the elasticity of the labor input factor:

$$\theta_{iLt} \equiv \left(\frac{\eta}{1+\eta}\right)\frac{W_{it}L_{it}}{P_{it}Q_{it}}(1-\mu_{it}^W). \tag{4.4}$$

Amoroso et al. (2012) define the bargained wage rate $\mu_{it}^W \equiv \frac{W_{it}-\bar{W}_{it}}{W_{it}}$ as the *wage markup*[29] From (4.4), after solving for L_{it} (see technical appendix), we derive the following expression for labor:

$$L_{it} = \left[(\exp(\theta_0+\omega_{it})K_{it}^{\theta_K}M_{it}^{\theta_M})^{\frac{\eta+1}{\eta}}\frac{1}{1-\mu^W}\frac{\eta+1}{\eta}\frac{\theta_{iLt}}{W_{it}}\frac{P_t^j}{(Q_t^j)^{1/\eta}}(\exp(u_{it}^d))^{-1/\eta}\right]^{\eta/(\eta-\theta_{iLt}(\eta-1))} \tag{4.5}$$

Substituting (4.5) into (4.3), taking into account (4.2), and assuming, for simplicity, that the elasticity of labor is constant across firms and time, we obtain the final short-run profit function:

$$\begin{aligned}&\Pi^{SR}(\omega_{it}, W_{it}, K_{it}, M_{it}, \psi_t) = \\ &\quad\left(\frac{1-\gamma}{\gamma^{1-\delta}}\right)W_{it}^{1-\delta}\left[\left(\exp(\theta_0+\omega_{it})K_{it}^{\theta_K}M_{it}^{\theta_M}\right)^{\frac{\eta+1}{\eta}}\left(\psi_t(\exp(u_{it}^d))^{-1/\eta}\right)\right]^{\delta}\end{aligned} \tag{4.6}$$

where $\psi_t \equiv \frac{P_t^j}{(Q_t^j)^{1/\eta}}$, $\gamma \equiv \theta_L\frac{\eta+1}{\eta}\frac{1}{1-\mu^W}$, and $\delta \equiv \eta/(\eta-\theta_{iLt}(\eta-1))$.

4.2.2 Dynamic decisions

The decisions of doing R&D, cooperating, or innovating cannot be revoked, so we assume the costs associated with these actions to be sunk. We define the vector of fixed costs paid in case of investment in research, cooperation, innovation, or both cooperation and innovating as $\theta_i^{FC} = (0, \theta_i^{FC}(rd), \theta_i^{FC}(c), \theta_i^{FC}(d), \theta_i^{FC}(cd))'$. We also define the vector of sunk costs associated with every investment choice k, $\theta_i^{SC} = (0, \theta_i^{SC}(rd), \theta_i^{SC}(c), \theta_i^{SC}(d), \theta_i^{SC}(cd))'$. In particular, we assume that, besides the fixed and sunk costs of R&D and innovation, there are sunk costs of finding an efficient research partner, or fixed costs of maintaining the research alliance, such as managing the contractual costs (transaction costs).

[29]In their paper, Amoroso et al. (2012) also show how, maximizing with respect to wages leads to an expression of the wage markup as a function of the bargaining parameter, ϕ_{it}, and the ratio between profits and cost of labor, $\mu_{it}^W = \frac{\phi_{it}}{1-\phi_{it}}\frac{\Pi_{it}}{W_{it}L_{it}}$.

Given their level of productivity, capital, materials, and present and past knowledge investment decisions, a_{it} and a_{it-1}, the firm faces the following profit function:

$$\begin{aligned}
&\Pi(a_{it}, a_{it-1}, \omega_{it}, W_{it}, K_{it}, M_{it}, \psi_t) = \\
&\quad \Pi^{SR}(\omega_{it}, W_{it}, K_{it}, M_{it}, \psi_t) - FC(K_{it}, M_{it}, a_{it}) - SC(a_{it}, a_{it-1}) \\
&\quad \equiv \Pi^{SR}(\omega_{it}, W_{it}, K_{it}, M_{it}, \psi_t) - \tilde{FC}(K_{it}, M_{it}) \\
&\quad -\theta_i'^{FC} a_{it} - \theta_i'^{SC}(1 - a_{it-1})a_{it}, \qquad (4.7)
\end{aligned}$$

where the function of the fixed costs of operation is defined as $FC(K_{it}, M_{it}, a_{it}) \equiv \tilde{FC}(K_{it}, M_{it}) - \theta_i'^{FC} a_{it-1}$

To simplify the framework, while retaining the salient features of the model, we make a set of assumptions. First, we omit the firm-level entry/exit decisions. Moreover, to reduce the dimensionality of the state vector on which firms are assumed to base their decisions, we consider a simpler framework, featuring imperfect competition only on the output market, and where capital and materials are assumed to be flexible inputs, not subject to adjustment costs. Assuming that the productivity, ω_{it}, and the aggregate state, ψ_t, are sufficient statistics for predicting the expected future profits, the short-run profit function under these restrictions is derived in the Appendix and can be written as

$$\Pi(a_{it}, a_{it-1}, \omega_{it}, \psi_t) = \varphi \psi_t \exp(\omega_{it})^{-(1+\eta)} - \theta_i'^{FC} a_{it} - \theta_i'^{SC}(1 - a_{it-1})a_{it}, \qquad (4.8)$$

where $\varphi \equiv -\frac{1}{1+\eta}\left(\frac{\eta}{1+\eta}\right)^{\eta}$.

4.2.2.1 State variables transition functions

We assume that the next period state of the aggregate variable ψ_t depends only on the current state. In particular, we specify the evolution of the aggregate state variable as

$$\psi_t = f(\psi_{t-1}) = \mu_0 + \rho\psi_{t-1} + \epsilon_\psi, \qquad (4.9)$$

where ϵ_ψ is a normally distributed error term. Following Santos (2009), the variance of ϵ_ψ, $\sigma_\epsilon^2 = \sigma_\psi^2(1 - \rho^2)$, represents the aggregate uncertainty of the industry affecting the firm's investment choice.

Concerning the productivity, we follow Doraszelski and Jaumandreu (2008), and Aw et al. (2011), and model the evolution of the firm's productivity as a Markov process, allowing for the productivity to be affected by firms' past choices of innovation, cooperation, and R&D.[30] We define the evolution process of productivity level ω_{it} of firm i at time t as:

$$\omega_{it} \equiv \omega(\omega_{it-1}, a_{it-1}) + \xi_{it} \tag{4.10}$$

where ξ_{it} is the normally distributed stochastic shock to productivity, and $\omega(\cdot)$ is approximated by a third degree polynomial.

In particular, we propose the evolution process of productivity level ω_{it} of firm i at time t as a nonlinearly persistent process, depending on a broader set of R&D activities, namely (cooperative) research and innovation. The productivity transition becomes:

$$\begin{aligned}\omega_{it} &= \omega(\omega_{it-1}, c_{it-1}, d_{it-1}, cd_{it-1}, rd_{it-1}) + \xi_{it} \qquad (4.11)\\ &= \beta_0 + \beta_1\omega_{it-1} + \beta_2\omega_{it-1}^2 + \beta_3\omega_{it-1}^3 + \beta_4 c_{it-1} + \beta_5 d_{it-1}\\ &\quad +\beta_6 c_{it-1}d_{it-1} + \beta_7 rd_{it-1} + \xi_{it}.\end{aligned}$$

The firm profit function as in (4.7) related to the set of choices a do not only differ in their fixed costs intercepts, but also in their arguments. In fact, the productivity process assumed in (4.11) depends on both the past level of productivity, and on the type of technological upgrade. Therefore, the variable ω_{it} associated with one choice might be different from that of an alternative investment choice.

Figure 4.1 reports the schematic representation of what the profile of all the optimal strategies for firm i and the relative payoffs, given the levels of productivity, could look like. The firms with a productivity level above a certain threshold decide to either invest in R&D ($\omega_{it} > \omega^{rd}$), or to cooperate with a research partner ($\omega_{it} > \omega^{c}$), as it might provide higher profits than doing R&D by themselves.

[30] Doraszelski and Jaumandreu (2008) relax the exogeneity assumption usually made about productivity in the production function literature (see Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg et al. (2006)), by letting the R&D spending and related activities to determine the differences in and the evolution of productivity across firms and over time. Aw et al. (2011) take a step further and assume that productivity evolves as a Markov process which depends on both investments in R&D and export market participation.

Cooperating yields higher profits since firms reduce the costs and associated risks of research by sharing them. Enterprises observing a level of productivity high enough to bear the sunk costs of introducing an innovation, invest in a product or process improvement that offers a greater performance or a reduced cost of production ($\omega_{it} > \omega^{d}$). Firms with productivity $\omega_{it} > \omega^{cd}$ engage in both activities and are thus assumed to be the most productive.

4.2.2.2 Value and policy functions

To retrieve information about the sunk costs of R&D, innovating, and cooperating, and to identify the evolution of the productivity states depending on firms' research investment policies, we consider a dynamic programming problem in which a firm i makes a series of discrete choices over its infinite lifetime.

Let a_{it} be the control variable and let S be the set of space state points and let the firms' characteristics s_{it} be an element of S. To simplify the framework, without losing the generality of the model, we assume that the state of firm i at time t is defined only by the level of productivity, ω_{it}, the industry competition proxied by the aggregate state ψ_t, and the past investment actions, a_{it-1}; therefore the state vector is summarized as $s_{it} = (\omega_{it}, \psi_t, a_{it-1})$. To fit the model to the data, we need to add unobserved heterogeneity. In particular, we introduce the vector of payoff shocks $\epsilon_{it} = \{\epsilon_{it}(k)\}_{k \in \{na,rd,c,d,cd\}}$ observed only by the firm. The unobserved characteristics ϵ_{it} are independently and identically distributed over time with continuous support and multivariate distribution function $F_{\epsilon}(\epsilon_{it})$. In particular, I assume that ϵ_{it}'s are *i.i.d.* extreme value distributed and enter the profit function in an additively separable way. These assumptions are not strictly necessary, but useful as they lead to a closed form likelihood function and a closed form expression for the expected maximum of the choice-specific value functions.

The observed state variable ω_{it} evolves as a Markov process depending stochastically on the choices of the firm because of the assumption in equation (4.10) with the cumulative distribution function given by $F_{\omega}(\omega_{it+1}|\omega_{it}, a_{it})$. On the other hand, the stochastic evolution of the aggregate state is assumed to be independent from the research activities, and therefore can be expressed as $F_{\psi}(\psi_{t+1}|\psi_t)$. Moreover, since we do not know the firm–level production technology, we assume the sunk costs of R&D, innovating, and of partnering in research to be drawn from a known joint distribution $F_{SC}(\theta_i^{SC})$.

Let us define $\theta_{\Pi i} \equiv ((\theta_i^{FC})', (\theta^{SC})')'$, and $\theta_\Pi \equiv \{\theta_{\Pi i}\}_{i=1,\ldots,N}$ as the matrix of choice- and firm-specific parameters that describe the profit function in (4.7). Finally, let $\theta = (vec(\theta_\Pi)', \theta'_\omega, \theta'_\psi, \theta'_\epsilon, \beta)' \in \Theta$ be the vector of the parameters of interest, where $vec(\theta_\Pi)$ is the vectorization of the θ_Π matrix, and where θ_ω and θ_ψ are vectors of parameters that describe the transition probability functions F_ω and F_ψ, respectively, θ_ϵ represents the parameters in the distribution of F_ϵ, and β is the rate at which the firm discounts future profits.

Assuming that firms behave optimally, the value function of firm i corresponds to the maximum of the expected discounted sum of profits, conditional on the current level of productivity and market indexes:

$$V(s_{it}, \epsilon_{it}; \theta) \equiv \max_{a_{it}, a_{it+1}, \ldots} E\left[\sum_{\tau=t}^{\infty} \beta^{\tau-t} \left(\Pi(a_{i\tau}, s_{i\tau}; \theta_{\Pi i}) + \epsilon_{i\tau}\right) | s_{it}, \epsilon_{it}\right] \qquad (4.12)$$

where $\beta \in (0,1)$, and $\Pi(a_{it}, s_{it}; \theta_{\Pi i}) + \epsilon_{it}$ are the current profits of firm i with productivity level ω_{it}, in market aggregate condition ψ_t, choosing investment a_{it}.

The problem is to determine, for all N firms, the set of optimal stationary decision rules $\alpha = \{\alpha_i\}_{i=1}^N$, where $\alpha_i : S \to A_i$, that solves the stochastic/multiperiod optimization problem expressed in (4.12). The method of dynamic programming offers the advantage of translating the optimization problem in (4.12) into a sequence of simpler deterministic/static optimization problems, where for $\beta \in (0,1)$ and for bounded $\Pi(\cdot)$, the value of the objective function can be written (suppressing the subscript i) in the form of a Bellman equation:

$$\begin{aligned} V(a, s, \epsilon; \theta) &= \Pi(a, s; \theta_\Pi) + \epsilon + \beta E_{s', \epsilon'}\left[V(s'; \theta)|s, a\right] \qquad (4.13) \\ V(s, \epsilon; \theta) &= \max_{a \in A} V(a, s, \epsilon; \theta) \end{aligned}$$

where s' and ϵ' denote the next period state and shock. Therefore, when conditioning on the value of the state and control variables, the optimal decisions of the firm do not depend on time t, but only on current and next period state variables. The assumption of the existence of a state variable that is designed to capture the productive and competitive environment faced by the firm at each point might be quite restrictive in the context of technological innovation. However, as in this paper, we consider the dynamic optimization problem of a single agent, the stationary dynamic programming framework could still capture the salient features of such a structural model.

The expected value function for next period is equal to:

$$E_{s',\epsilon'}\left[V(s',\epsilon';\theta)|s,a\right] = \int_{s'}\int_{\epsilon'} V(s',\epsilon';\theta)dF_{\epsilon}(\epsilon';\theta_{\epsilon})dF_s(s'|s,a;\theta), \tag{4.14}$$

where $dF_s(s'|s,a;\theta) \equiv dF_{\omega}(\omega'|\omega,a;\theta_{\omega})dF_{\psi}(\psi'|\psi;\theta_{\psi})$. Given that the optimal strategy, $\alpha(s,\epsilon)$, satisfies

$$\alpha(s,\epsilon) = \arg\max_{a\in A} V(a,s,\epsilon;\theta),$$

and observing data $(\mathbf{a},\boldsymbol{\omega},\boldsymbol{\psi}) \equiv \{\{a_{it},\omega_{it}\}_{i=1}^{N},\psi_t\}_{t=1}^{T}$, in order to estimate θ, we construct the likelihood as the product of firms' conditional choice probabilities (CCPs), $P_{it}(a_{it}|s_{it};\theta)$, as

$$\begin{aligned} P_{it}(a_{it}|s_{it};\theta) &\equiv Pr(\epsilon : V(a_{it},s_{it};\theta) \geq V(\tilde{a}_{it},s_{it};\theta)), \quad \forall \tilde{a}_{it} \\ &= Pr(\epsilon : a_{it} = \alpha(s_{it},\epsilon_{it})) \\ &= \int \mathbb{1}\{a_{it} = \alpha(s_{it},\epsilon_{it})\}dF_{\epsilon}. \end{aligned}$$

The joint likelihood of the observed data is then:

$$L(\mathbf{a}|\mathbf{s};\theta) = \prod_i\prod_t P_{it}(a_{it}|s_{it};\theta). \tag{4.15}$$

Moreover, since ϵ follows a joint Gumbel (extreme value type I) distribution, independent across alternatives k, the likelihood increment for firm i is

$$P_{it}(a_{it}|s_{it};\theta) = \frac{\exp\{V(\tilde{a}_{it},s_{it};\theta)\}}{\sum_{a_{it}\neq\tilde{a}_{it}}\exp\{V(a_{it},s_{it};\theta)\}}. \tag{4.16}$$

In the next section, we discuss the empirical strategy to estimate the static structural parameters, namely, the demand elasticity, the wage markup, the aggregate state proxying the industry competitive environment, the productivity evolution parameters, the fixed costs, and the dynamic parameters, i.e., the sunk costs, and the discount factor.

4.3 The estimation procedure

Estimation is done in three steps. In the first step, we estimate a production function that allows us to retrieve estimates of the firm-level productivity, ω_{it}, the parameters describing the aggregate state and productivity evolution processes, $f(\psi_{t-1})$, and $\omega(\omega_{it-1}, a_{it-1})$, repectively, and the structural parameters needed to construct the profit function as in (4.6). In the second step, we retrieve the management costs concerning the research activity adopted by the firm. In the last step, we obtain estimates of the dynamic structural parameters, $\theta_\Pi, \theta_\omega, \theta_\psi, \theta_\epsilon$, by numerical approximation of the solution to the dynamic programming problem at trial parameters.

4.3.1 Step 1: Static parameters

The production function and the demand parameters are estimated with the method proposed by Amoroso et al. (2012). Within the Cobb-Douglas production function framework, they relax the conventional assumption of perfect competition in the labor market, allowing both firms and workers' union to have some market power.

In their study, Amoroso et al. (2012) report empirical evidence of the underestimation of the true level of price–cost margins caused by the omission of direct effects of the wage bill on marginal costs. In fact, the exclusion of frictions in the labor market (i.e., $\phi_{it} = 0$ or $W_{it} = \bar{W}_{it}$) might lead to misestimating the firm's market power. When there is no imperfect competition in the labor market, firms set the wage at the lowest value possible, ultimately equal to the competitive wage, i.e., $W_{it} = \bar{W}_{it}$ (and, therefore, $\mu_{it}^W = 0$). For W_{it} that tends to $\bar{W}_{it}$, the wage markup decreases, given that the elasticity and the share of labor are constant, which is inversely related to the output markup $\frac{\eta}{1+\eta}$.

Next to the labor market rigidities, Amoroso et al. (2012) also correct for the possible bias in the estimated coefficients when deflated gross output is used instead of gross physical output. Defining the log deflated output as y_{it}, this can be rewritten as

$$y_{it} = q_{it} + (p_{it} - p_t^j),$$

where p_t^j is the log industry price index. The firm-level price deviations $(p_{it} - p_t^j)$ will enter the production function as an extra error component, introducing potential correlation with the input choices. Substituting p_{it} with the inverse Dixit-Stiglitz demand function, and taking into account the labor input elasticity under imperfect competition in the labor market,

$$\theta_{iLt}\left(\frac{\eta+1}{\eta}\right) \equiv \gamma_{iLt} = s_{iLt}(1-\mu_{it}^W), \tag{4.17}$$

where s_{iLt} is the share of labor and it is defined as the ratio between cost of labor and total sales $\left(\frac{W_{it}L_{it}}{P_{it}(Q_{it})Q_{it}}\right)$, they estimate a log deflated revenue function that features both labor and output market distortions:

$$y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + (1-\mu_{it}^W)s_{iLt}l_{it} - \frac{1}{\eta}q_t^j + \tilde{\omega}_{it} + \tilde{u}_{it} \tag{4.18}$$

where k_{it}, l_{it}, m_{it} are logs of deflated capital, labor, and deflated materials, respectively; q_t^j is the log of the production index in sector j. The composite error term, $\tilde{u}_{it} \equiv u_{it}^q + u_{it}^d$, includes the demand shock, $\tilde{u}_{it}^d \equiv -u_{it}^d/\eta$, and the measurement error, u_{it}^q. $\tilde{\omega}_{it} \equiv \omega_{it}(1+\eta)/\eta$ is the productivity.

The production index is constructed as in De Loecker (2011), by proxying the total demand for a sector j with a (market share) weighted average of deflated revenue, $q_t^j = \sum_i^{N_j} ms_{it}y_{it}$. Both the intercept, $\gamma_0 \equiv \theta_0(1+\eta)/\eta$, and the factor elasticities of capital and material, $\gamma_k \equiv \theta_k(1+\eta)/\eta, k = K, M$ are divided by the *output price markup* defined as $\equiv \eta/(1+\eta)$ for $\eta < -1$. The elasticity of labor is defined as in (4.17).

The firm–level productivity ω_{it} is estimated as

$$\hat{\omega}_{it} = \hat{\eta}/(1+\hat{\eta})\tilde{\omega}_{it} = \hat{\eta}/(1+\hat{\eta})\left[y_{it} - \left(\hat{\gamma}_0 + \hat{\gamma}_K k_{it} + \hat{\gamma}_M m_{it} + (1-\hat{\mu}_{it}^W)s_{iLt}l_{it} - \frac{1}{\hat{\eta}}q_t^j\right)\right].$$

Identification of all the structural parameter of the deflated revenue function in (4.18) is ensured by the presence of firm specific wages. To estimate all the relevant parameters, they adopt a control function approach (Olley and Pakes, 1996) which consists in including additional regressors to capture the endogenous part of the unobserved productivity. In particular, the productivity $\tilde{\omega}_{it}$ can be approximated by a third–degree polynomial (Levinsohn and Petrin, 2003) in all three factor inputs k_{it}, l_{it}, m_{it}. The productivity is also assumed to evolve over time as a Markov process that depends on the firms' investment choices, as in (4.11). The replacement function approach allows for dynamics in the productivity process, but restricts the investment function, and consequently the productivity process, to be homogeneous across firms. On the other hand, the instrumental variables approach comes at the cost of not allowing for the possibility that the unobserved productivity could be correlated with past choices of inputs. Therefore, for the problem at hand, we rely on the control function approach to identify the deflated revenue function parameter, and our object of interest, the firm level productivity. The estimation of (4.18) requires the following moment restrictions

$$E(\xi_{it} + \tilde{u}_{it} | m_{it}, k_{it}, l_{it-1}, m_{it-1}, k_{it-1}, \ldots, l_{i1}, m_{i1}, k_{i1}) = 0,$$

however, identification could hold with just current values and one lag in the conditioning set.

Results of the estimation of the deflated revenue function under imperfect competition in both output and labor markets (4.18), of the aggregate state transition function (4.9), and of the nonlinearly persistent productivity process depending on technology upgrading (4.11) are reported and discussed in Section 4. In the following subsection, we discuss the second step of our estimation strategy, namely, how to retrieve the fixed costs of (cooperative) research and innovation.

4.3.2 Step 2: Profit function parameters

It is well-known that, in general, the parameters of structural dynamic programming problems are not identified (Rust, 1994). Magnac and Thesmar (2002) show that the utility functions of the firms can be identified if the distribution function of the unobserved preference shocks, the discount rate, and the value function of one the alternatives (normalization) are fixed. Hence, it is theoretically possible to identify both fixed and sunk costs of R&D and innovation. However, in practice the simultaneous identification of such costs requires enough variation in the observed R&D investment decisions. To circumvent this problem, we recover the fixed cost parameters within the static framework, after having estimated the production function parameters. In particular, we consider the estimation of the fixed costs of innovative investments as a random utility model (multinomial mixed logit model), where the alternative-specific utility function of firm i is associated with the level of productivity and fixed costs represent the alternative-specific firm-level random coefficients associated with the research ivestment k, i.e.,

$$V(a_{it}, s_{it}, \zeta_{it}; \theta^{FC}) = \varphi\psi_t^{\eta} \exp(\omega_{it})^{-(1+\eta)} - \theta_i'^{FC} a_{it} + \zeta_{it}.$$

The error term ζ_{it} is a random term assumed to be iid extreme value distributed. To identify θ_i^{FC}, we assume that the additive separable utility shock ζ_{it} is exogenous. Results of this estimation are reported in Section 5.

4.3.3 Step 3: Dynamic parameters

The main limiting factor in estimating dynamic discrete choice (DDC) models is the computational complexity resulting form the need to compute the continuation values as in (4.14). The direct way of obtaining such continuation values has been to compute them as the fixed point of a functional equation. For example, Rust (1987) proposes a computational strategy named the nested fixed point (NFXP) algorithm, which is a gradient iterative search method to obtain the maximum likelihood estimator of the structural parameters. Unfortunately, the NFXP algorithm is computationally demanding because it requires to obtain the fixed point of a Bellman operator (hence, it must run successive iterations of the value functions until convergence) for each point in the state space of the structural parameters. Additionally, the number of state points grows exponentially with the dimensionality of the state space. This concern about the computational burden of implementing the NFXP algorithm, and the curse of dimensionality, have led to a number of estimators that are computationally faster (Bajari et al., 2007; Pakes et al., 2007). For example, the two-step estimator by Hotz and Miller (1993), using nonparametric estimates of choice and state transition probabilities, yields a simple representation of the choice-specific value functions for values in a neighborhood of the true vector of structural parameters.[31] The main advantage of this two-step estimator is its computational simplicity. The first step is a nonparametric regression to obtain the productivity and the aggregate state transition functions, the second step is the estimation of a standard discrete choice model (the policy functions) with a criterion function that is globally concave (e.g., such as the likelihood of a multinomial logit model in our investment choice study case). Thus, the agent's continuation values can be obtained nonparametrically by first estimating the agent's choice probabilities at each state, and then inverting the choice problem to obtain the corresponding continuation values. However, as with other approaches, there are limitations. First, since the two-step empirical strategy involves the (nonparametric) estimation of the CCPs, the continuation values are estimated rather than computed, and therefore they contain sampling error. This sampling error might be significant if the state space of the model is large relative to the available data. The second limitation comes from the formal requirements of the limit properties of the estimator. As a matter of fact, to obtain an estimator with desirable properties, the data must visit a subset of the points repeatedly. More precisely, all the states in some recurrent class $\Re \subseteq S$ must be visited infinitely often, and the equilibrium strategies must be the same every time each point of $\Re$ is visited. Simply put, the two-step approach requires the assumption of stationarity. To give an example, when forecasting the CCPs of a firm observed in year t when being active on the market in year $t+\tau$, it is assumed that the firm at time t would face the same decision-making environment observed in year $t+\tau$. Moreover, it must also be assumed that there is no permanent unobserved heterogeneity, otherwise, it would be impossible to match the actions of the firm at time t with the action at time $t+\tau$.

To correct for the finite sample bias, Aguirregabiria and Mira (2002) propose a nested pseudo-likelihood algorithm (NPL) for the estimation of the class of discrete Markov decision models with the conditional independence assumption.

[31]For an exhaustive, but self-contained review and description of Hotz and Miller (1993) two-step estimator and extensions, see Aguirregabiria and Mira (2010).

In particular, their method considers a K-step extension of the Hotz and Miller (1993) estimator. In fact, Aguirregabiria and Mira (2002) obtain a new estimate of the CCPs given the two-step estimator and an initial nonparametric estimator of the CCPs. Successive iterations return a sequence of estimators of the structural parameters and CCPs that are asymptotically equivalent to the partial MLE and to the two-step PML (Aguirregabiria and Mira, 2002, Proposition 4). Moreover, Aguirregabiria and Mira (2002) report results from Monte Carlo experiments that illustrate how iterating in this procedure does in fact produce significant reductions in finite sample bias. However, their estimation algorithm have difficulties dealing with unobserved heterogeneity. Extensions to accommodate unobserved heterogeneity via finite mixture distributions into CCP estimation are attributable to Arcidiacono and Miller (2011).

Given these recent extensions, there is still one main limiting factor in estimating DP models, which is the computational burden associated with the iterative process. Therefore, it is not surprising that there have been continuing efforts to reduce the computational burden of estimating DP models. Recently, computationally practical Bayesian approaches that rely on Markov Chain Monte Carlo (MCMC) methods have been developed by Imai et al. (2009) and Norets (2009).

In this paper, we adopt the estimation method proposed by Imai et al. (2009). Their algorithm is related to the one proposed by Aguirregabiria and Mira (2002), but it is based on the full solution of the DP problem, yielding the advantage of dealing with unobserved heterogeneity. The main idea of their estimation approach is to avoid the computation of the full solution of the DP problem, by approximating the expected value function at a state space point using the average of value functions at past iterations in which the parameter vector is close to the current parameter vector and the state variables are close to the current state variables.[32] In the conventional NFXP algorithm, most of the information

[32] Ching et al. (2012) claim that the practical Bayesian approach developed by Imai et al. (2009)

> "...is potentially superior to prior methods because (1) it could significantly reduce the computational burden of solving for the DDP model in each iteration, and (2) it produces the posterior distribution of parameter vectors, and the corresponding solutions for the DDP model–this avoids the need to search for the global maximum of a complicated likelihood function."

obtained in the past iterations remains unused in the current iteration.

The Imai et al. (2009) algorithm consists of two loops:

1. **The outer loop (Metropolis-Hasting Algorithm)**
   The outer loop performs a M-H (Metropolis-Hasting) algorithm. First, we draw a candidate parameter vector from a proposal density, then we evaluate the likelihood, conditional on the candidate parameter vector and on the previous iteration parameter vector, to compute the acceptance probability, with which we can decide whether or not to accept the candidate parameter vector.

   In our setting, we allow for the parameters of the profit function, θ_Π, to take different values for each firm. In particular, we assume that the vector of firm-specific parameters $\theta_{\Pi i}$ follows the density function:

   $$\theta_{\Pi i} \sim g(\theta_{\Pi i}(a); \mu),$$

   where $\mu = (\bar{\theta}_\Pi, \sigma_\Pi)'$ is the hyperparameter vector for this density. In particular, we assume g is a normal distribution and μ includes parameters for means, $\bar{\theta}_\Pi$, and standard deviations, σ_Π. Assuming that the prior of the mean parameters is normal and that of the standard deviation parameters is inverted Gamma, the posterior distribution for the mean parameters is normal and that for the standard deviation parameter is inverted Gamma. To simplify the framework, without losing the generality of the structural model, we assume that the priors are independent across investment alternatives.

   The entire parameter vector consists now of $\theta = (\mu', vec(\theta_\Pi)', \theta'_\omega, \theta'_\psi, \theta'_\epsilon, \beta)'$. Let us rewrite this vector as $\theta = (\mu', vec(\theta_\Pi)', (\theta_c)')'$, where $\theta_c = (\theta'_\omega, \theta'_\psi, \theta'_\epsilon, \beta)'$ is the vector of parameters common across firms. As for the prior on θ_c, we use independent flat priors. Suppose we are at iteration r, with parameter estimates being $(\mu^r, vec(\theta_\Pi), \theta_c)$, then the outer loop iteration for drawing a parameter vector from the posterior distribution can be divided into three steps:

   1.1 *Hyperparameter updating step*

Draw μ^r. That is, given θ_Π^{r-1}, for all alternative $a \in A$, draw $\bar{\theta}_\Pi \sim f_\theta(\cdot|\sigma_{\theta_\Pi}^{r-1}, \{\theta_{\Pi i}^{r-1}\}_{i=1}^N)$ and $\sigma_{\Pi(a)}^r \sim f_\sigma(\cdot|\bar{\theta}_\Pi^r, \{\theta_{\Pi i}^{r-1}\}_{i=1}^N)$, where f_θ and f_σ are the conditional posterior distributions.

1.2 *Data augmentation step*

Now that we have effectively constructed the prior for $\theta_{\Pi i}$, we draw, for each alternative a, a candidate parameter from the proposal density, which we assume to be a normal density,

$$\theta_{\Pi i}^{*r} \sim q(\bar{\theta}_\Pi^{r-1}, \sigma_{\theta_\Pi}^{r-1}).$$

Then, accept $\theta_{\Pi i}^{*r}$ with probability λ, where

$$\lambda = \min\left\{\frac{g(\theta_{\Pi i}^{*r};\mu^r)P_i^r(a_i|\omega_i,\psi;\theta_{\Pi i}^{*r},\theta_c^{r-1})q(\cdot|\theta_{\Pi i}^{*r},\mu^r)}{g(\theta_{\Pi i}^{r-1};\mu^r)P_i^r(a_i|\omega_i,\psi;\theta_{\Pi i}^{r-1},\theta_c^{r-1})q(\cdot|\theta_{\Pi i}^{r-1},\mu^r)}, 1\right\}.$$

The computation of the firm-specific likelihood component P_i^r, as defined in (4.16), requires the computation of the expected value function for the firm, which happens in the inner loop.

1.3 *Common parameters drawing step*

We draw a candidate parameter form the proposal density $\theta_c^{*r} \sim q(\theta_c^{*r}|\theta_c^{r-1})$, then accept θ_c^{*r} with probability λ, where

$$\lambda = \min\left\{\frac{\pi(\theta_c^{*r})L^r(\mathbf{a}|\boldsymbol{\omega},\psi;\theta_\Pi^r,\theta_c^{*r})q(\cdot|\theta_c^{*r})}{\pi(\theta_c^{r-1})L^r(\mathbf{a}|\boldsymbol{\omega},\psi;\theta_\Pi^r,\theta_c^{r-1})q(\cdot|\theta_c^{r-1})}, 1\right\},$$

where $(\mathbf{a},\boldsymbol{\omega}) \equiv \{a_i,\omega_i\}_{i=1}^N$, and L^r is the joint likelihood defined in (4.15).

2. **The inner loop**

The inner loop computes and updates the alternative specific value function by applying the Bellman operator once. Imai et al. (2009) propose to approximate the expected value functions by storing and using information from earlier iterations of the algorithm. In particular, storing up to M past accepted draws of parameters and value functions, $\{\theta^{*l}, s^l, V^l(s^l, \epsilon^l; \theta^{*l})\}_{l=r-M}^{r-1}$, Imai et al. (2009) propose to construct the expected value function in iteration r as,

$$E^r_{\epsilon'}\left[V(s', \epsilon'; \theta^{*r}|s, a)\right] = \sum_{l=r-M}^{r-1} V^l(s^l, \epsilon^l; \theta^{*l})\chi(\theta^{*l}, \theta^{*r}; s^l, s|a), \tag{4.19}$$

where

$$\chi(\theta^{*l}, \theta^{*r}; s^l, s|a) = \frac{K_{h_\theta}(\theta^{*l}, \theta^{*r})K_{h_s}(s^l, s|a)}{\sum_{k=r-M}^{r-1} K_{h_\theta}(\theta^{*k}, \theta^{*k})K_{h_s}(s^k, s|a)},$$

so as to assign higher weights to past parameters that are closer the current iteration one, and higher weights to states s' that have higher transition density from states s. $K_{h_\theta}(\theta^{*k}, \theta^{*k})$ and $K_{h_s}(s^k, s|a)$ are kernel function with bandwidth h_θ, and h_s, for the parameter vector, θ, and the state variable s, respectively. The value function obtained from (4.19) is used to construct the choice specific value function,

$$V^r(a, s, \epsilon; \theta^{*r}) = \Pi(a, s; \theta^{*r}_\Pi) + \epsilon + \beta E^r_{\epsilon'}\left[V(s', \epsilon'; \theta^{*r})|s, a\right]. \tag{4.20}$$

The value function in (4.20) is used to construct the likelihood as in (4.16). Note that the integration over the continuous state variables is already incorporated into the computation of the weighted average of past value functions. This approach has the advantage, compared to Rust's random grid approximation, of avoiding to compute the value function at N_{grid} random points of the state variables state in each iteration.

Finally, given the assumption of *iid* extreme value distributed ϵ's, we have that

$$V^r(s, \epsilon; \theta^{*r}) = \max_{a\in A} V(a, s, \epsilon; \theta^{*r}) = \ln\left[\sum_a \exp(V(a, s; \theta^{*r}))\right].$$

4.4 Data

In this section, we report the summary statistics of all the variables used to estimate the static and the dynamic structural models. In particular, the upper part of Table 4.2 displays mean, standard deviation, and number of observation of the variables extracted from the PS (Production Survey, Statistics Netherlands) for the years 2002-2008. To estimate the deflated revenue function as in (4.18), we use the deflated value of gross output Y_{it} ($\equiv \frac{P_{it}Q_{it}}{\tilde{P}_t^j}$) of each firm i in sector j in period t, where $P_{it}Q_{it}$ are the firm's revenues, and $\tilde{P}_t^j$ is the sector j price deflator. Labor (L_{it}) refers to the number of employees in each firm for each year,[33] collected in September of that year. The corresponding wages W_{it} include gross wages plus salaries and social contributions before taxes. The costs of intermediate inputs ($Z_{it}M_{it}$) include costs of energy, intermediate materials, and services. The unit user costs R_{it} (of capital stock K_{it}) are calculated as the sum of the depreciation of fixed assets and the interest charges. Q_t^j indicates the sector-specific production index.

The nominal gross output and intermediate inputs are deflated with the appropriate price indices from the input-output tables available at the NACE rev. 1 two-digits sector classification.[34] For capital, we use a two-digit NACE deflator of fixed tangible assets calculated by Statistics Netherlands. The share of the cost of labor, material, and capital are denoted as s_{iLt}, s_{iMt}, and s_{iKt}, respectively. The share of the cost of labor constitutes 24.2 percent of the gross production value, while materials account for 65.7 percent of gross output, and capital for 4 percent.

The total number of observation, after retaining only the respondents to the different waves of the Community Innovation Survey, is 8306. The CIS datasets are the main data source for measuring innovation in Europe. The surveys are designed to provide an extensive description of the general structure of innovative activities at the sectoral, regional, and country levels, including basic information

[33]For each enterprise, jobs are added and adjusted for part-time and duration factors, resulting in number of men/years expressed as Full Time Equivalents (FTEs)(*Source:* Statistics Netherlands)

[34]NACE Rev. 1 is a 2-digit activity classification which was drawn up in 1989. It is a revision of the General Industrial Classification of Economic Activities within the European Communities, known by the acronym NACE and originally published by Eurostat in 1970.

of the enterprise, product and process innovation, innovation activity and expenditure, effects of innovation, innovation cooperation, public finding of innovation, source of information for innovation patents, and so forth.[35]

The middle part of Table 4.2 reports descriptive statistics for the different types of R&D expenditure extracted from three waves of the Community Innovation Survey (CIS), carried out by Statistics Netherlands. In particular, we constructed an unbalanced panel of survey respondents, merging the CIS 4 (reference period 2002-2004), the CIS 2006 (reference period 2004-2006), and the CIS 2008 surveys (2006-2008). The R&D expenditures are expressed in thousands of Euros. The intramural expenditure are more than three times larger than the extramural. The average total amount of research expenditure is roughly 3 million Euros. The number of firms that reported R&D spending is 2171 out the total sample of 3565 (unevenly distributed over the period 2002-2008). The last part of Table 4.2 displays the details of the control variable, namely the investment choice k. The most right column reports the total number of firms for each year. For example, in 2002, the number of enterprises that participated to the CIS and that were matched with the PS is 444 , whereas in 2008, the same matching exercise yields a much larger number of firms, i.e., 2413. Our R&D investment variable is constructed as follows. The firm-specific choice na_{it} takes value one if the firm does not engage in any activity other than operating in the market; rd_{it} takes value one if the firm decides to spend in R&D; the investment decision c_{it} takes value of one if the firm has at least one cooperative agreement (with either a firm, a supplier, a customer, or a public (private) research institute); d_{it} match firms' decision to invest in a technological upgrade; action cd_{it} tags the decision to both innovate and cooperate. Concerning the type of investment, the simple production without innovative or cooperative activities is the most frequent, with a total of 3389 observations ($k = na$). Introducing an innovation (product or process, $k = d$), and both innovating and cooperating with either another firm ($k = cd$), or with a research institute are also very frequent answers (2129 and 2530 observation, respectively). On the other hand, the number of firms engaging in only R&D ($k = rd$) or only research alliances ($k = c$) is quite small, with an average of 23 and 13 firms for the rd and c investment choices, respectively.

[35] Community Innovation Survey, EUROSTAT.

The cross-sectional data from each wave is expanded so as to cover the whole reference period (there is a one-year overlap between the three waves). For example, if the firm has reported to have introduced an innovation during the reference period, and the same firm has not abandoned the innovation project, then we impute the value 1 for the whole time span.

4.5 Results

In this section we first present the parameter estimates of the deflated revenue function under imperfect competition in both output and labor markets, (4.18), and of the state variables evolution, (4.9), and (4.11). We then use the estimates of the static parameters to present the results of the dynamic discrete choice model.

4.5.1 Static parameters

The point estimates of the output price markup and all the parameters used to construct the productivity evolution as in (4.11) are reported in Table 4.1. The upper part of the table reports demand elasticity parameters, the aggregate state average, and the productivity level and growth.[36] The elasticity of the demand is found to be equal to -2.8, with a corresponding output price markup of 55%. On average, the log productivity is equal to 1.381 and its growth is equal to 1.7%. The aggregate state, ψ_t, is constructed as weighted deflated total industry revenues, $\psi_t \equiv \sum_j \tilde{p}_t^j / N^j (q_t^j)^{1/\eta}$, where $\tilde{p}_t^j$ is the price deflator for industry j at time t, and q_t^j is the weighted average of deflated revenues per industry. We find the aggregate state to be equal to 1.088, on average. Analyzing the evolution of the aggregate state over the years, we find that the market conditions were stable until 2006 and start worsening in 2007 and 2008. The same pattern is followed by the total factor productivity (TFP) growth. The correlation between ψ_t and productivity is 0.922 (significant at 0.001 significance level). These results confirm

[36]For a complete discussion on the factor input elasticities and the implication of the rent-sharing parameter on productivity growth, we would refer the reader to the paper of Amoroso et al. (2012).

that, at an aggregate level, the TFP growth estimated under the assumptions of imperfect competition in both labor and output markets seems to pick up the actual features of the Dutch manufacturing industry.

The aggregate state transition of (4.9) is specified by the three estimated parameters, the mean, $\hat{\mu}_0 = 0.853$, the autocorrelation, $\hat{\rho} = 0.241$, and the variance, $\hat{\sigma}_\epsilon = 0.114$.

Concerning the parameters of the productivity evolution as in (4.11), we find evidence of a third order polynomial, and fair dependence on innovation and cooperation. In particular, the estimated coefficient associated with the action of cooperating is significant at the 5% significance level, and equal to 0.076, and that of innovating is equal to 0.113. The coefficient associated with both cooperating and upgrading technology, and the decision to do R&D, are equal to 0.062 and -0.011, respectively.

The four means and standard errors of the posterior distributions of the fixed costs are reported in Table 4.3. Assuming that all firms face the same log-normal distribution for all four fixed costs, we find that the fixed costs of R&D and cooperating in R&D (3.0 and 3.5 million Euro, respectively) are substantially higher than the per-period costs of maintaining an innovation (460 thousand Euro). Moreover, the fixed costs of maintaining an innovative activity, while sharing the costs of R&D, decreases the per-period costs (290 thousand Euro). This confirms the rationale behind the cooperating strategies, i.e., the cost sharing motive (Cassiman and Veugelers, 2002; Lopez, 2008; Amoroso, 2011). R&D cooperations, in fact, allow firms to share costs or to reduce risks of innovation. The results for the fixed costs are comparable with those found by Aw et al. (2011) for the Taiwanese electronics industry, as they estimate these costs to be on average 67.606 million TW dollars (roughly 1.8 million Euro)

Below the posterior means and standard deviation of the fixed costs relative to each innovative activity, we report the probabilities of undertaking the different investments, given the level of productivity and the market conditions. On average, the probability to not engage in any activity is the highest (0.41), followed by the probability to simultaneously cooperate and innovate (0.30), and by the probability to introduce an innovation (0.26). Next to the averages of the probability of choosing action k, we report the same probabilities for the levels of the log productivity at each quartile. As we are interested in understanding the relation between the level of productivity and the probability of undertaking an activity, Figure 4.2 displays the locally weighted scatterplot smoothing (lowess)[37] curves fitting the relationships between the probabilities to undertake action a and the level of productivity, $\exp(\omega_{it})$. The darker areas of the smoothed scatter-plots represent higher density of the data points. The plot at the top reports the curve fitting the relation between the probability of taking no action and the level of productivity. The probability of remaining inactive in research and innovation is inversely related to the productivity. We find the same pattern for the probability of doing R&D and the probability of introducing an innovation. Simply put, the higher the firm level productivity, the smaller the probability of investing in R&D, or innovating. However, the situation is reversed when the investment in R&D or in a new product or process is shared with a partner. Indeed, when cooperating, the probabilities of doing research, $Pr(a = c|s, \theta)$, and innovating, $Pr(a = cd|s, \theta)$, are (non monotonic) increasing functions of productivity. This pattern could point to the presence of knowledge externalities. These results, together with the evidence of the endogenous firm-level productivity, positively associated with the action of cooperating, suggest that an innovation policy aiming at encouraging research cooperation might result in a virtuous cycle. Indeed, past investments in cooperative research have a positive impact on current productivity, which, in turn, positively influence the probability to engage in both R&D and innovation when these activities are shared with a research partner. Figure 4.3 plots the MCMC draws of the fixed cost parameters. It appears that the the MCMC draws converge after 50 iterations.

4.5.2 Dynamic parameters

In this section we present the results for the DDP model presented in (4.13). Once the fixed costs are estimated, we can subtract them from the profit function as in (4.7). For simplicity, we estimate the model without unobserved heterogeneity. Therefore, the standard deviations σ_Π are set equal to zero. The discount factor

[37]Locally weighted regression fitting techniques provide a generally smooth curve, the value of which at a particular location along the x-axis is determined only by the points in that vicinity. The method consequently makes no assumptions about the form of the relationship, and allows the form to be discovered using the data itself.

is fixed at 0.93. During this stage we are able to recover both fixed and sunk costs of doing R&D or innovating with or without a research partner. Figure 4.3 shows that the sunk cost parameters converge at different rates, and, in general, much slower than the fixed costs.

The estimated coefficients are reported in Table 4.4. Next to the mean values of the sunk costs, we report the standard deviations of the MCMC draws. The values are estimated with the expected signs. Sunk costs are found to be 4 millions for the average firm that undertakes R&D with or without a partner, 14 to 33% higher than the fixed costs. The sunk costs of innovation are still much smaller than the ones of research, but 3 to 3.5 times higher than the fixed costs of innovating. Moreover, we find additional evidence of the risk-sharing motive behind the decision to introduce an innovation. In fact, the average sunk costs of producing an innovation with a research partner is almost one third smaller than the average sunk costs of undertaking the same project without an alliance (997,000 Euro and 1.4 million Euro, respectively). The sunk costs parameters cannot be compared with the reported R&D expenditures. This is because the sunk costs can be related to productive factors, such as labor and/or capital that are allocated to research rather than to production. For this reason, these costs will not appear in the balance sheets of the company (Santos, 2009).

Next to the estimation of the sunk cost parameters, we show the importance of the role played by these costs in shaping the probabilities if undertaking the different research investments. Table 4.4 also reports the changes in probabilities associated with 50% and 25% reductions in the costs of engaging in research and/or innovating. A reduction in the sunk costs of R&D, cooperating, and innovating can be thought of as an example of an innovation policy, such as a subsidy to R&D start up, or public procurement. Results show that a 25% reduction in these costs is expected to increase the probability of undertaking the corresponding activity. For example, reducing the costs of R&D, $\theta_i^{SC}(rd)$ of 25% leads to an increase of probability of doing R&D of 62.2%.

4.6 Conclusion

In this paper, we present empirical evidence of the fixed and sunk costs of investments in research activities, and quantify the linkages between the cost structure, firm-level productivity, and the probabilities to technologically upgrade. In particular, we propose and estimate a structural model with endogenous choices of technological upgrade for the Dutch manufacturing industry. The model describes a firm's dynamic decision process for undertaking different research activities, namely, innovating and conducting R&D, with or without a research partner. The R&D investment choices are endogenous, as they depend on the firm's level of productivity, an aggregate measure of industry competition, fixed and sunk costs of R&D, and past research choices. To our knowledge, none of the existing studies proposes and estimates a dynamic structural model to derive the total cost function of firms engaging in technological activities.

We find that the firm's probability to do R&D or to introduce an innovation increases with the level of productivity, only when this activity is shared with a research partner. Moreover, according to the literature on R&D cooperation, the costs of innovating are smaller when cooperating. In fact, given the higher risks associated with the uncertainty of the market demand for new products or processes, the firm might allocate more importance to the cost/risk sharing rationale for this type of innovative activities, rather than for the sheer research investments.

Sunk costs are found to be roughly 1.5 times larger than the fixed costs of research (both cooperative and private), and 3 to 3.5 times larger than the fixed costs of innovating. Moreover, we show the importance of the role played by these costs in shaping the probabilities if undertaking the different research investments. In general, a reduction in the sunk costs of R&D, cooperating, and innovating increases the probability of undertaking the corresponding activity.

Additionally, we present some preliminary conclusions on innovation policies aiming at encouraging research cooperation. We show how these type of policy interventions might result in a virtuous cycle. Indeed, past investments in cooperative research have a positive impact on current productivity, which, in turn, positively influences the probability to engage in both R&D and innovation when these activities are shared with a research partner. Therefore, in elaborating their policies for innovation, governments must ensure to create frameworks that encourage the collaboration throughout the innovation process.

4.7 Appendix

4.7.1 Profit function

Given the following maximization problem

$$\max_{L_{it},W_{it}} \left[\phi_{it}\log(U_{it}(W_{it},L_{it})) + (1-\phi_{it})\log\Pi_{it}\right],$$

the first order conditions can be written as:

$$w.r.t. \quad L_{it} \quad \rightarrow \quad (1-\phi_{it})\frac{W_{it} - \left(1+\frac{1}{\eta}\right)P_{it}(Q_{it})\frac{\partial Q_{it}}{\partial L_{it}}}{\Pi_{it}} = \frac{\phi_{it}}{L_{it}}, \tag{4.21}$$

$$w.r.t. \quad W_{it} \quad \rightarrow \quad (1-\phi_{it})\frac{W_{it} - \bar{W}_{it}}{\Pi_{it}} = \frac{\phi_{it}}{L_{it}}. \tag{4.22}$$

Combining equations (4.21) and (4.22), the marginal revenue product of labor is

$$\left(\frac{\eta+1}{\eta}\right)P_{it}(Q_{it})\frac{\partial Q_{it}}{\partial L_{it}} = \bar{W}_{it}. \tag{4.23}$$

Therefore, by multiplying both sides of (4.23) by $\frac{L_{it}}{Q_{it}}$, we have

$$\frac{\eta+1}{\eta}\theta_{iLt} = \frac{\bar{W}_{it}L_{it}}{P_{it}(Q_{it})Q_{it}} = \frac{\bar{W}_{it}}{W_{it}}\frac{W_{it}L_{it}}{P_{it}(Q_{it})Q_{it}}.$$

Using Amoroso et al. (2012) definition of the wage markup $\mu^W_{it} \equiv \frac{W_{it}-\bar{W}_{it}}{W_{it}}$, and taking into account the demand as in (4.1), we can rewrite the cost of labor as

$$W_{it}L_{it} = \frac{1+\eta}{\eta}\theta_{iLt}\frac{1}{1-\mu^W_{it}}(Q_{it})^{\frac{1+\eta}{\eta}}\frac{P^j_t}{(Q^j_t)^{1/\eta}}\exp(-u^d_{it}/\eta).$$

Replacing Q_{it} with the Cobb-Douglas function as in (4.2), and solving for L_{it}, we get

$$L_{it} = \left[(\exp(\theta_0+\omega_{it})K^{\theta_K}_{it}M^{\theta_M}_{it})^{\frac{\eta+1}{\eta}}\frac{1}{1-\mu^W}\frac{\eta+1}{\eta}\frac{\theta_{iLt}}{W_{it}}\frac{P^j_t}{(Q^j_t)^{1/\eta}}(\exp(-u^d_{it}/\eta))\right]^{\eta/(\eta-\theta_{iLt}(\eta-1))}. \tag{4.24}$$

The short-run profits, $P_{it}Q_{it} - W_{it}Lit$, can be rewritten as

$$\Pi^{SR}(\omega_{it}, W_{it}, K_{it}, M_{it}, \psi_t) =$$
$$(\exp(\theta_0 + \omega_{it})K_{it}^{\theta_K} L_{it}^{\theta_L} M_{it}^{\theta_M})^{\frac{1+\eta}{\eta}} \frac{P_t^j}{(Q_t^j)^{1/\eta}} \exp(-u_{it}^d/\eta) \left[1 - \frac{1+\eta}{\eta}\theta_{iLt}\frac{1}{1-\mu_{it}^W}\right].$$

Replacing the labor demand with (4.24), we get the final profit function:

$$\Pi^{SR}(\omega_{it}, W_{it}, K_{it}, M_{it}, \psi_t) =$$
$$\left(\frac{1-\gamma}{\gamma^{1-\delta}}\right) W_{it}^{1-\delta} \left[\left(\exp(\theta_0 + \omega_{it})K_{it}^{\theta_K} M_{it}^{\theta_M}\right)^{\frac{\eta+1}{\eta}} \left(\psi_t(\exp(u_{it}^d))^{-1/\eta}\right)\right]^{\delta}$$

where $\psi_t \equiv \frac{P_t^j}{(Q_t^j)^{1/\eta}}$, $\gamma \equiv \theta_L \frac{\eta+1}{\eta}\frac{1}{1-\mu^W}$, and $\delta \equiv \eta/(\eta - \theta_{iLt}(\eta - 1))$.

The short-run profit function as in (4.8), assuming no imperfect competition on the labor market, is derived from the following optimization problem for firm i:

$$\max_{X_{it}} \ \{P_{it}Q_{it} - V_{it}'X_{it} \mid A_{it}F(X_{it}) \geq Q_{it}\}, \tag{4.25}$$

where $X_{it} \equiv (X_{i1t}, X_{i2t}, \ldots, X_{irt})'$ denotes the vector of r factor inputs, $F(.)$ is production function, and $V_{it} \equiv (V_{i1t}, V_{i2t}, \ldots, V_{irt})'$ is the vector of r input prices. Taking into account the demand as in (4.1), the FOC is:

$$\frac{\eta+1}{\eta}P_{it}\frac{\partial Q_{it}}{\partial X_{it}} = V_{it},$$

since $MC_{it}^X = V_{it}\frac{\partial X_{it}}{\partial Q_{it}}$ is defined as the marginal cost of X_{it}, we have that

$$\frac{P_{it} - MC_{it}^X}{P_{it}} = -\frac{1}{\eta}. \tag{4.26}$$

Assuming that the marginal cost of X_{it} are an inverse function of the firm-level productivity such as

$$MC_{it}^X \equiv \frac{1}{\exp(\omega_{it})},$$

the price can be expressed as a function of the demand elasticity and the productivity,

$$P_{it} = \frac{\eta}{\eta+1}\frac{1}{\exp(\omega_{it})}. \tag{4.27}$$

Multiplying (4.26) by $P_{it}Q_{it}$, we obtain the profits, therefore the profit function can be written as

$$\Pi_{it} = -\frac{1}{\eta} P_{it} Q_{it}.$$

Substituting Q_{it} with (4.1) and P_{it} with (4.27), we obtain the following short-run profit function:

$$\Pi(\omega_{it}, \psi_t) = \varphi \psi_t \exp(\omega_{it})^{-(1+\eta)},$$

where $\varphi \equiv -\frac{1}{1+\eta} \left(\frac{\eta}{1+\eta}\right)^{\eta}$.

4.7.2 Tables and Figures

Table 4.1: Demand and productivity evolution parameters

	parameter	estimate	(st.err.)/st.dvt.
	θ_L	0.266	(0.036)
	θ_M	1.206	(0.114)
Eq. (4.18)	θ_K	0.044	(0.010)
	η	-2.800	(0.428)
	$\eta/(\eta+1)$	1.555	(0.132)
	μ^W	0.311	(0.050)
	φ	0.332	(0.000)
	ψ_t	1.088	0.178
	ω_{it}	1.381	0.327
	$\Delta\omega_{it}$	0.017	0.225
	μ_0	0.853	(0.022)
Eq. (4.9)	ρ	0.241	(0.020)
	σ_ϵ	0.114	(0.001)
	β_0	1.650	(0.020)
	β_1	0.581	(0.043)
	β_2	-0.002	(0.002)
	β_3	0.001	(0.000)
Eq. (4.11)	β_4	0.076	(0.043)
	β_5	0.113	(0.205)
	β_6	0.062	(0.056)
	β_7	-0.011	(0.077)

Table 4.2: Summary Statistics

	mean	sd	median	1^{st} quartile	3^{rd} quartile	N. obs
$P_{it}Q_{it}$	63323.97(K Euros)	318679	14881.500	5838.000	39925	8306
L_{it}	152.657	347.055	75	36	152	8306
$Z_{it}M_{it}$	48353.050(K Euros)	280848	9868	3539	27120	8306
$R_{it}K_{it}$	2255.667(K Euros)	26330	359	117	1156	8257
s_{iLt}	0.242	0.124	0.228	0.154	0.310	
s_{iMt}	0.657	0.149	0.663	0.567	0.758	
s_{iKt}	0.040	0.223	0.027	0.013	0.048	
Q_t^j	73.080	10.465	73.498	63.648	80.889	8306
Intramural R&D	1806.574(K Euros)	18396.654	100	10	400	4937
Extramural R&D	612.855(K Euros)	7243.232	0	0	50	4937
R&D Expenditure	3038.461(K Euros)	26356.650	255	63	846	4937
	$k = na$	$k = rd$	$k = c$	$k = d$	$k = cd$	N_t
N_{2002}	153	22	9	136	124	444
N_{2003}	133	9	7	102	167	418
N_{2004}	175	13	6	131	221	546
N_{2005}	179	8	3	130	184	504
N_{2006}	769	28	15	471	557	1840
N_{2007}	907	38	26	553	617	2141
N_{2008}	1073	46	28	606	660	2413
Tot.	3389	164	94	2129	2530	8306

Table 4.3: Fixed costs

	posterior mean($\times 1mln$)		std error	
$\theta_i^{FC}(rd)$	3.025		0.082	
$\theta_i^{FC}(c)$	3.528		0.100	
$\theta_i^{FC}(d)$	0.459		0.025	
$\theta_i^{FC}(cd)$	0.286		0.025	
	mean	$\omega_{it} \leq 1.172$	$\omega_{it} \leq 1.347$	$\omega_{it} \leq 1.541$
$P_i(a = na\|s, \theta)$	0.408	0.409	0.415	0.414
$P_i(a = rd\|s, \theta)$	0.020	0.024	0.021	0.021
$P_i(a = c\|s, \theta)$	0.011	0.008	0.009	0.011
$P_i(a = d\|s, \theta)$	0.256	0.269	0.263	0.260
$P_i(a = cd\|s, \theta)$	0.304	0.288	0.289	0.293

Figure 4.1: R&D, Cooperation and Innovation Choices

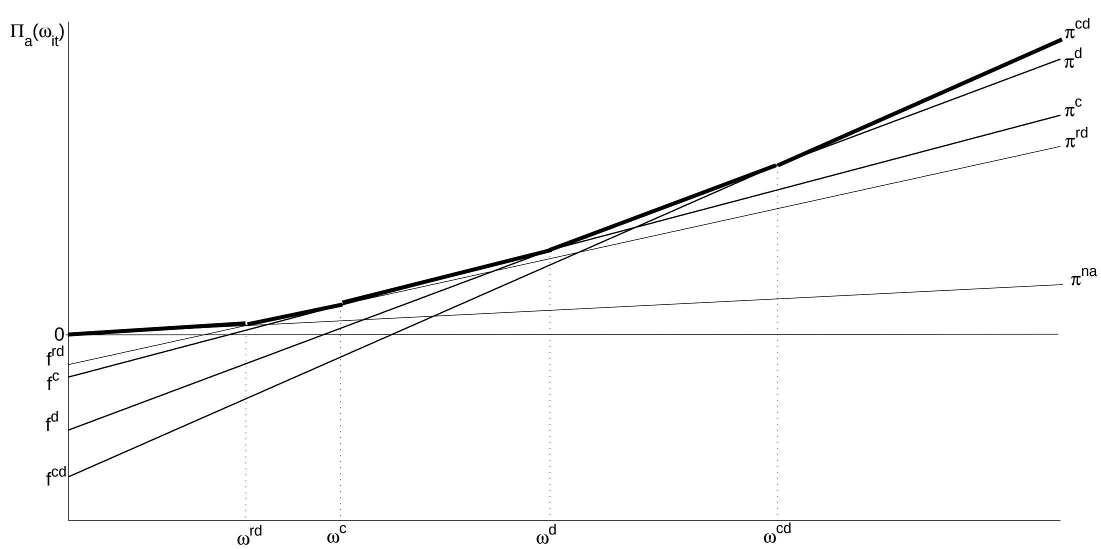


Figure 4.2: Investment policy functions

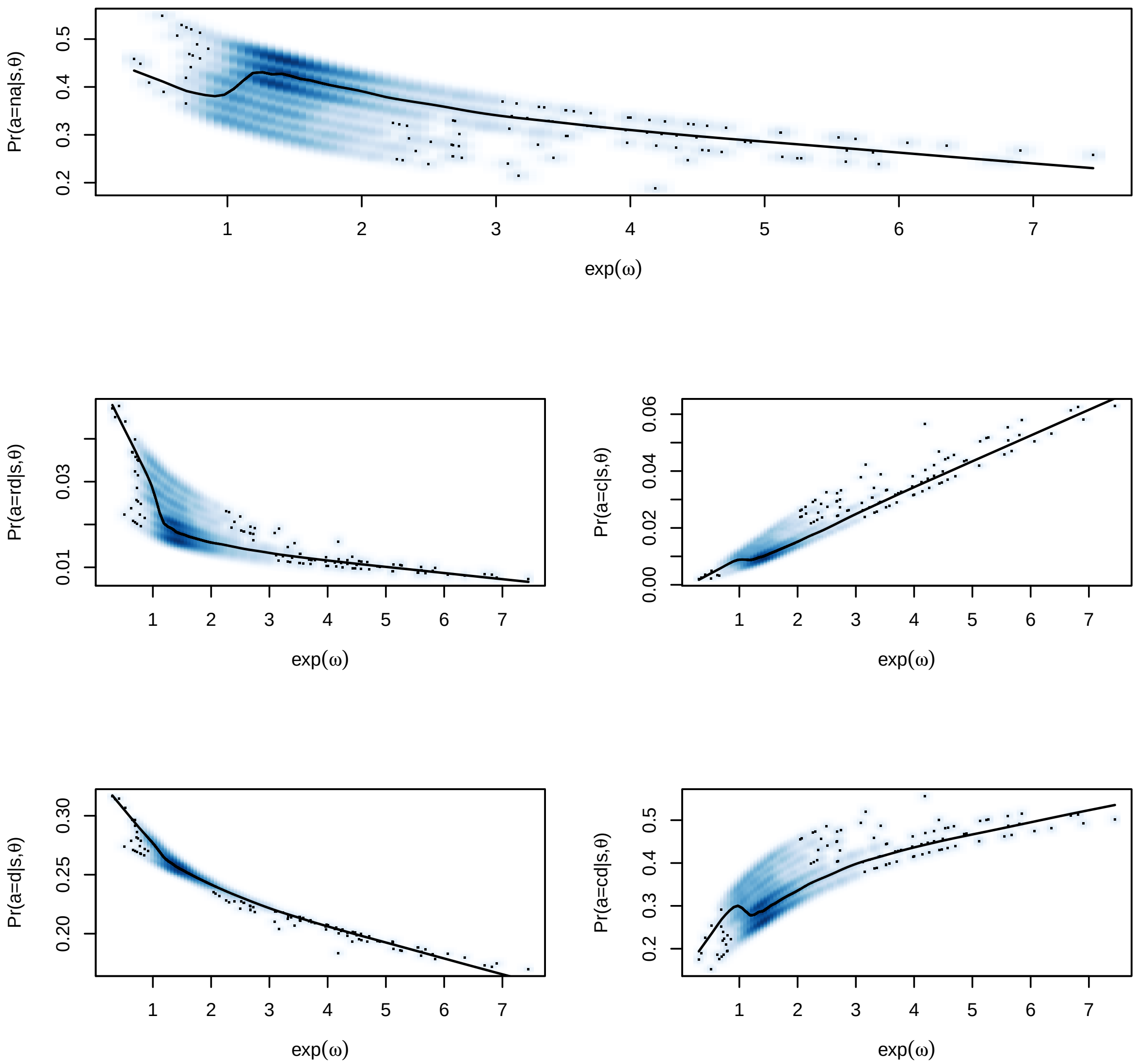
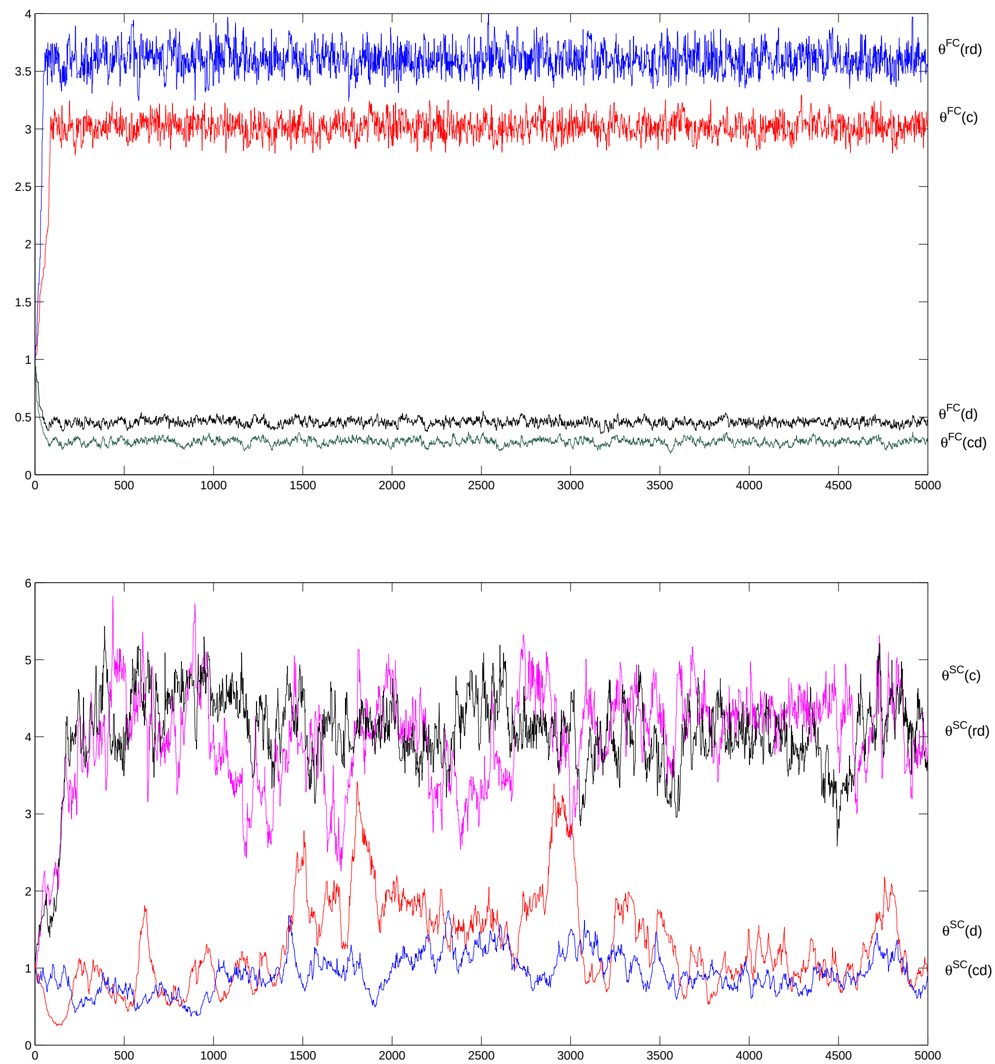


Table 4.4: Dynamic Parameter Estimates

	posterior mean		std error
$\theta_i^{SC}(rd)$	3.984		0.570
$\theta_i^{SC}(c)$	4.046		0.216
$\theta_i^{SC}(d)$	1.433		0.560
$\theta_i^{SC}(cd)$	0.997		0.216

$\theta_i^{SC}(rd)$	-50%	-25%	0%
$P_i(a = na\|s, \theta)$	0.141	0.264	0.367
$P_i(a = rd\|s, \theta)$	0.049	0.630	0.008
$P_i(a = c\|s, \theta)$	0.037	0.006	0.009
$P_i(a = d\|s, \theta)$	0.002	0.048	0.390
$P_i(a = cd\|s, \theta)$	0.770	0.051	0.226
$\theta_i^{SC}(c)$	-50%	-25%	0%
$P_i(a = na\|s, \theta)$	0.484	0.693	”
$P_i(a = rd\|s, \theta)$	0.006	0.002	”
$P_i(a = c\|s, \theta)$	0.065	0.008	”
$P_i(a = d\|s, \theta)$	0.263	0.109	”
$P_i(a = cd\|s, \theta)$	0.182	0.188	”
$\theta_i^{SC}(d)$	-50%	-25%	0%
$P_i(a = na\|s, \theta)$	0.668	0.433	”
$P_i(a = rd\|s, \theta)$	0.002	0.005	”
$P_i(a = c\|s, \theta)$	0.002	0.007	”
$P_i(a = d\|s, \theta)$	0.225	0.407	”
$P_i(a = cd\|s, \theta)$	0.103	0.147	”
$\theta_i^{SC}(cd)$	-50%	-25%	0%
$P_i(a = na\|s, \theta)$	0.616	0.541	”
$P_i(a = rd\|s, \theta)$	0.007	0.003	”
$P_i(a = c\|s, \theta)$	0.004	0.004	”
$P_i(a = d\|s, \theta)$	0.129	0.162	”
$P_i(a = cd\|s, \theta)$	0.244	0.290	”

Figure 4.3: MCMC iterations of fixed and sunk cost parameters



Note: MCMC plots of θ^{FC} and θ^{SC}

References

Abraham, F., Konings, J., Vanormelingen, S., 2009. The effect of globalization on union bargaining and price–cost margins of firms. Review of World Economics 145, 13–36.

Ackerberg, D., Benkard, L., Berry, S., Pakes, A., 2007. Econometric Tools for Analyzing Market Outcomes. Handbook of Econometrics 6.

Ackerberg, D., Caves, K., Frazer, G., 2006. Structural Identification of Production Functions. Tech. rep., UCLA mimeo.

Aguirregabiria, V., Mira, P., 2002. Swapping the nested fixed point algorithm: A class of estimators for discrete markov decision models. Econometrica 70 (4), pp. 1519–1543.

Aguirregabiria, V., Mira, P., May 2010. Dynamic discrete choice structural models: A survey. Journal of Econometrics 156 (1), 38–67.

Aidt, T. S., Tzannatos, Z., 2008. Trade unions, collective bargaining and macroeconomic performance: a review. Industrial Relations Journal 39 (4), 258–295.

Almus, M., Czarnitzki, D., April 2003. The effects of public r&d subsidies on firms' innovation activities: The case of eastern germany. Journal of Business & Economic Statistics 21 (2), 226–36.

Amoroso, S., 2011. Bayesian analysis of r&d cooperation determinants. Working Paper.

Amoroso, S., Melenberg, B., Plasmans, J., Vancauteren, M., 2012. Firm level productivity under imperfect competition in output and labor markets. Working Paper.

Arcidiacono, P., Miller, R. A., 2011. Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. Econometrica 79 (6), 1823–1867.

Arellano, M., Bond, S., April 1991. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. Review of Economic Studies 58 (2), 277–97.

Aw, B. Y., Roberts, M. J., Xu, D. Y., 2011. R&d investment, exporting, and productivity dynamics. American Economic Review 101 (4), 1312–44.

Bajari, P., Benkard, C. L., Levin, J., 09 2007. Estimating dynamic models of imperfect competition. Econometrica 75 (5), 1331–1370.

Belderbos, R., Carree, M., Diederen, B., Lokshin, B., Veugelers, R., 2004a. Heterogeneity in R&D cooperation strategies. International Journal of Industrial Organization 22 (8-9), 1237–1263.

Belderbos, R., Carree, M., Lokshin, B., 2004b. Cooperative R&D and firm performance. Research Policy 33 (10), 1477–1492.

Belderbos, R., Carree, M., Lokshin, B., 2006. Complementarity in R&D cooperation strategies. Review of Industrial Organization 28 (4), 401–426.

Blundell, R., Bond, S., 2000. Gmm estimation with persistent panel data: an application to production functions. Econometric Reviews 19 (3), 321–340.

Bond, S., Söderbom, M., Jan. 2005. Adjustment costs and the identification of cobb douglas production functions. Economics Papers 2005-W04, Economics Group, Nuffield College, University of Oxford.

Boone, J., 2008. A new way to measure competition. The Economic Journal 118 (531), 1245–1261.

Boone, J., van der Wiel, H., 2007. How (not) to measure competition. CPB Discussion Paper 91, CPB Netherlands Bureau for Economic Policy Analysis.

Bryan, L. L., 2007. The new metrics of corporate performance : profit per employee. The McKinsey quarterly (1), 57–65.

Bughin, J., 1993. Union-firm efficient bargaining and test of oligopolistic conduct. The Review of Economics and Statistics 75 (3), 563–567.

Bughin, J., 1996. Trade unions and firms' product market power. The Journal of Industrial Economics 44 (3), pp. 289–307.

Busom, I., Fernández-Ribas, A., 2008. The impact of firm participation in R&D programmes on R&D partnerships. Research Policy 37 (2), 240–257.

Carboni, O. A., 2012. An empirical investigation of the determinants of r&d cooperation: An application of the inverse hyperbolic sine transformation. Research in Economics 66 (2), 131 – 141.

Cassiman, B., Veugelers, R., Sep 2002. R&d cooperation and spillovers: Some empirical evidence from belgium. Open access publications from katholieke universiteit leuven, Katholieke Universiteit Leuven.

Cassiman, B., Veugelers, R., Jan 2006. In search of complementarity in innovation strategy: Internal r&d and external knowledge acquisition. Open access publications from katholieke universiteit leuven, Katholieke Universiteit Leuven.

Ching, A. T., Imai, S., Ishihara, M., Jain, N., 2012. A practitioner's guide to bayesian estimation of discrete choice dynamic programming models. Quantitative Marketing and Economics 10, 151–196.

Cohen, W. M., Levinthal, D. A., 1990. Absorptive capacity: A new perspective on learning and innovation. Administrative Science Quarterly 35 (1), pp. 128–152.

Coull, B., Agresti, A., 2000. Random effects modeling of multiple binomial responses using the multivariate binomial logit-normal distribution. Biometrics 56 (1), 73–80.

Crépon, B., Desplatz, R., Mairesse, J., 2002. Price-cost margins and rent sharing: Evidence from a panel of french manufacturing firms. Revised version of CREST Working Paper No. G9917.

Crépon, B., Duguet, E., Mairessec, J., 1998. Research, innovation and productivity: An econometric analysis at the firm level. Economics of Innovation and New Technology 7 (2), 115–158.

d'Aspremont, C., Jacquemin, A., December 1988. Cooperative and noncooperative r&d in duopoly with spillovers. American Economic Review 78 (5), 1133–37.

De Loecker, J., 2011. Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. Econometrica 79 (5), 1407–1451.

De Loecker, J., Konings, J., June 2006. Job reallocation and productivity growth in a post-socialist economy: Evidence from slovenian manufacturing. European Journal of Political Economy 22 (2), 388–408.

Dewar, R. D., Dutton, J. E., 1986. The adoption of radical and incremental innovations: an empirical analysis. Management science 32 (11), 1422–1433.

Dobbelaere, S., 2004. Estimation of price-cost margins and union bargaining power for belgian manufacturing. International Journal of Industrial Organization 22, 1381–1398.

Dobbelaere, S., Mairesse, J., 2011. Panel data estimates of the production function and product and labor market imperfections. Journal of Applied Econometrics.

Doraszelski, U., Jaumandreu, J., Jan. 2008. R&d and productivity: Estimating production functions when productivity is endogenous. CEPR Discussion Papers 6636, C.E.P.R. Discussion Papers.

Dosi, G., 1999. Innovation Policy in a Global Economy. Cambridge University Press, Ch. Some Notes on National Systems of Innovation and Production, and their Implications for Economic Analysis.

Ericson, R., Pakes, A., 1995. Markov-perfect industry dynamics: A framework for empirical work. The Review of Economic Studies 62 (1), pp. 53–82.

Foster, L., Haltiwanger, J., Syverson, C., March 2008. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? American Economic Review 98 (1), 394–425.

Galí, J., Gertler, M., Lpez-Salido, J. D., November 2007. Markups, gaps, and the welfare costs of business fluctuations. The Review of Economics and Statistics 89 (1), 44–59.

Gandhi, A., Navarro, S., Rivers, D., 2011. On the identification of production functions: How heterogeneous is productivity? University of Western Ontario, CIBC Centre for Human Capital and Productivity Working Papers 20119, University of Western Ontario, CIBC Centre for Human Capital and Productivity.

Gelman, A., Carlin, J., Stern, H., Rubin, D., 2003. Bayesian data analysis, 2nd Edition. Chapman and Hall.

Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to the calculationof posterior moments. In: Berger, J., Bernardo, J., Dawid, A., Smith, A. (Eds.), Bayesian Statistics. Oxford University Press, pp. 169–194.

Goldstein, H., 1995. Multilevel Statistical Models, 2nd Edition. Halstead Press, New York.

Griffith, R., Boone, J., Harrison, R., 2008. Measuring competition. Social Science Research Network Working Paper Series.

Griliches, Z., 1980. R & d and the productivity slowdown. The American Economic Review 70 (2), pp. 343–348.

Grimpe, C., Kaiser, U., 2010. Balancing internal and external knowledge acquisition: The gains and pains from r&d outsourcing. Journal of Management Studies 47 (8), 1483–1509.

Hadfield, J., Kruuk, L., 2010. MCMC methods for multi-response generalised linear mixed models: The MCMCglmm R package. Journal of Statistical Software 33 (2), 1–22.

Hall, B. H., Mairesse, J., January 1995. Exploring the relationship between r&d and productivity in french manufacturing firms. Journal of Econometrics 65 (1), 263–293.

Hall, R. E., 1986. Market structure and macroeconomic fluctuations. Brookings Papers on Economic Activity 17 (2), 285–338.

Hall, R. E., October 1988. The relation between price and marginal cost in u.s. industry. Journal of Political Economy 96 (5), 921–47.

Hall, R. E., October 1991. Invariance properties of solow's productivity residual. Working Paper 3034, National Bureau of Economic Research.

Heckman, J. J., Lochner, L., Taber, C., May 1998. General-equilibrium treatment effects: A study of tuition policy. American Economic Review 88 (2), 381–86.

Hedeker, D., Gibbons, R. D., 1996. Mixor: a computer program for mixed-effects ordinal regression analysis. Computer Methods and Programs in Biomedicine 49, 157–176.

Heidelberger, P., Welch, P. D., 1983. Simulation run length control in the presence of an initial transient. Operations Research 31 (6), 1109–1144.

Henderson, R., 1993. Underinvestment and incompetence as responses to radical innovation: evidence from the photolithographic alignment equipment industry. RAND Journal of Economics 24 (2), 248–270.

Hernán, R., Marín, P. L., Siotis, G., 03 2003. An empirical evaluation of the determinants of research joint venture formation. Journal of Industrial Economics 51 (1), 75–89.

Hotz, V. J., Miller, R. A., 1993. Conditional choice probabilities and the estimation of dynamic models. The Review of Economic Studies 60 (3), pp. 497–529.

Howells, J., 1999. Research and technology outsourcing. Technology Analysis & Strategic Management 11 (1), 17–29.

Imai, S., Jain, N., Ching, A., November 2009. Bayesian estimation of dynamic discrete choice models. Econometrica 77 (6), 1865–1899.

Jones, C. I., Williams, J. C., November 1998. Measuring the social return to r&d. The Quarterly Journal of Economics 113 (4), 1119–1135.

Kaiser, U., 2002. An empirical test of models explaining research expenditures and research cooperation: evidence for the german service sector. International Journal of Industrial Organization 20 (6), 747 – 774.

Kamien, M. I., Muller, E., Zang, I., December 1992. Research joint ventures and r&d cartels. American Economic Review 82 (5), 1293–306.

Katayama, H., Lu, S., Tybout, J. R., May 2009. Firm-level productivity studies: Illusions and a solution. International Journal of Industrial Organization 27 (3), 403–413.

Katz, M. L., 1986. An analysis of cooperative research and development. RAND Journal of Economics 14 (4), 527–543.

Kirat, T., Lung, Y., 1999. Innovation and Proximity. European Urban and Regional Studies 6 (1), 27–38.

Klette, T. J., Griliches, Z., July-Aug. 1996. The inconsistency of common scale estimators when output prices are unobserved and endogenous. Journal of Applied Econometrics 11 (4), 343–61.

Klette, T. J., Moen, J., Griliches, Z., April 2000. Do subsidies to commercial r&d reduce market failures? microeconometric evaluation studies1. Research Policy 29 (4-5), 471–495.

Korotayev, A. V., Tsirel, S. V., 2010. A spectral analysis of world gdp dynamics: Kondratieff waves, kuznets swings, juglar and kitchin cycles in global economic development, and the 20082009 economic crisis. Structure and Dynamics 4 (1).

Leifer, R., Gina Colarelli, O., Rice, M., Gina Colarelli, O., 2001. Implementing radical innovation in mature firms: The role of hubs. The Academy of Management Executive (1993-2005) 15 (3), 102–113.

Levinsohn, J., Petrin, A., 04 2003. Estimating production functions using inputs to control for unobservables. Review of Economic Studies 70 (2), 317–341.

Lopez, A., January 2008. Determinants of r&d cooperation: Evidence from spanish manufacturing firms. International Journal of Industrial Organization 26 (1), 113–136.

Magnac, T., Thesmar, D., March 2002. Identifying dynamic discrete decision processes. Econometrica 70 (2), 801–816.

Masayuki, Morikawa, 2010. Labor unions and productivity: An empirical analysis using japanese firm-level data. Labour Economics 17 (6), 1030 – 1037.

McDonald, I. M., Solow, R. M., December 1981. Wage bargaining and employment. American Economic Review 71 (5), 896–908.

McDonald, I. M., Suen, A., May 1992. On the measurement and determination of trade union power. Oxford Bulletin of Economics and Statistics 54 (2), 209–24.

Mohnen, P., Roller, L.-H., August 2005. Complementarities in innovation policy. European Economic Review 49 (6), 1431–1450.

Norets, A., 2009. Inference in dynamic discrete choice models with serially orrelated unobserved state variables. Econometrica 77 (5), 1665–1682.

OECD, E., 1997. Proposed Guidelines for Collecting and Interpreting Technological Innovation Data: Oslo Manual.

Olley, G. S., Pakes, A., 1996. The dynamics of productivity in the telecommunications equipment industry. Econometrica 64 (6), pp. 1263–1297.

Ornaghi, C., 2008. Price deflators and the estimation of the production function. Economics Letters 99 (1), 168 – 171.

Pakes, A., Ostrovsky, M., Berry, S., 2007. Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). The RAND Journal of Economics 38 (2), 373–399.

Ramey, V. A., Shapiro, M. D., 2001. Displaced capital: A study of aerospace plant closings. Journal of Political Economy 109 (5), pp. 958–992.

Reinganum, J., 1983. Uncertain innovation and the persistence of monopoly. The American Economic Review, 741–748.

Rodríguez, G., Goldman, N., 1995. An assessment of estimation procedures for multilevel models with binary responses. J. Royal Statistical Society, 73–90.

Rust, J., September 1987. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. Econometrica 55 (5), 999–1033.

Rust, J., 1994. Chapter 51 structural estimation of markov decision processes. In: Engle, R. F., McFadden, D. L. (Eds.), Handbook of Econometrics. Vol. 4 of Handbook of Econometrics. Elsevier, pp. 3081 – 3143.

Santos, C. D., Nov. 2009. Recovering the sunk costs of r&d: the moulds industry case. CEP Discussion Papers dp0958, Centre for Economic Performance, LSE.

Schmitz, H., 1999. Collective efficiency and increasing returns. Cambridge Journal of Economics 23 (4), 465–83.

Tether, B., 2002. Who co-operates for innovation, and why:: An empirical analysis. Research Policy 31 (6), 947–967.

Train, K., 2009. Discrete Choice Methods with Simulation, 2nd Edition. Online economics textbooks. Cambridge University Press.

Veugelers, R., 1997. Internal R&D expenditures and external technology sourcing. Research policy 26 (3), 303–315.

Wang, L., Zajac, E., 2007. Alliance or acquisition? A dyadic perspective on interfirm resource combinations. Strategic Management Journal 28 (13), 1291–1317.

Wang, X. H., Yang, B. Z., 2001. Fixed and sunk costs revisited. Journal of Economic Education 32 (2), 178–185.

Wooldridge, J. M., September 2009. On estimating firm-level production functions using proxy variables to control for unobservables. Economics Letters 104 (3), 112–114.

Zeger, S., Karim, M., 1991. Generalized linear models with random effects; a Gibbs sampling approach. Journal of the American statistical association 86 (413), 79–86.